TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 370, Number 1, January 2018, Pages 577–599
http://dx.doi.org/10.1090/tran/6971
Article electronically published on July 19, 2017

INTRINSIC DIOPHANTINE APPROXIMATION ON MANIFOLDS: GENERAL THEORY

LIOR FISHMAN, DMITRY KLEINBOCK, KEITH MERRILL, AND DAVID SIMMONS

Abstract. We investigate the question of how well points on a nondegenerate k-dimensional submanifold $M \subseteq \mathbb{R}^d$ can be approximated by rationals also lying on M, establishing an upper bound on the "intrinsic Dirichlet exponent" for M. We show that relative to this exponent, the set of badly intrinsically approximable points is of full dimension and the set of very well intrinsically approximable points is of zero measure. Our bound on the intrinsic Dirichlet exponent is phrased in terms of an explicit function of k and d which does not seem to have appeared in the literature previously. It is shown to be optimal for several particular cases. The requirement that the rationals lie on M distinguishes this question from the more common context of (ambient) Diophantine approximation on manifolds, and necessitates the development of new techniques. Our main tool is an analogue of the simplex lemma for rationals lying on M which provides new insights on the local distribution of rational points on nondegenerate manifolds.

1. Introduction and motivation

In its classical form, the field of Diophantine approximation addresses questions regarding how well points $\mathbf{x} \in \mathbb{R}^d$ can be approximated by rational points, where the quality $\|\mathbf{x} - \mathbf{r}\|$ of a rational approximation $\mathbf{r} \in \mathbb{Q}^d$ is compared with the size of the denominator of $\mathbf{r}$. The most fundamental theorem in the field is Dirichlet's theorem, dating back to 1842, which establishes a rate of approximation which holds for *every* point.[1] The full significance of this result was realized two years later when Liouville established that certain real numbers, namely quadratic irrationals, do not admit a better rate of approximation. Liouville's result was later generalized to higher dimensions by Perron [27]; together, these results show that Dirichlet's theorem is optimal in every dimension, in a sense to be made rigorous below.

Questions related to Diophantine approximation can be asked in a much broader context. Consider a closed subset M of a complete metric space (X, dist), a countable subset $\mathcal{Q} \subseteq X$ and a *height* function $H\colon \mathcal{Q} \to (0,\infty)$. Modifying the terminology recently introduced in [13], we will refer to such a collection as a *Diophantine triple*, and denote it by $(M, \mathcal{Q}, H)$. Given such a triple, we can then look for a function $\psi\colon (0,\infty) \to (0,\infty)$ such that for every $\mathbf{x} \in M$, there exists a constant $C_{\mathbf{x}}$

Received by the editors September 29, 2015 and, in revised form, April 13, 2016.
2010 *Mathematics Subject Classification.* Primary 11J13, 11J83; Secondary 37A17.
The first-named author was supported in part by the Simons Foundation grant #245708. The second-named author was supported in part by the NSF grant DMS-1101320. The fourth-named author was supported in part by the EPSRC Programme Grant EP/J018260/1.

[1]It was recently pointed out to us by Bugeaud that in the case of $\mathbb{R}$, this result is actually much older, coming directly from the theory of continued fractions (see e.g. [24, displayed equation on p. 28]).

and a sequence $\mathbf{r}_n \in \mathcal{Q}$ $(n \in \mathbb{N})$ satisfying

$$\mathbf{r}_n \to \mathbf{x} \quad \text{and} \quad \operatorname{dist}(\mathbf{x}, \mathbf{r}_n) \leq C_{\mathbf{x}} \psi\big(H(\mathbf{r}_n)\big) \text{ for all } n.$$

We will call such a function ψ a *Dirichlet* function for approximation of points in M by $\mathcal{Q}$. (Note that M has to be contained in the closure of $\mathcal{Q}$ in order for a Dirichlet function to exist.) Using this terminology, and defining the height $H\colon \mathbb{Q}^d \to (0,\infty)$ as $H(\mathbf{p}/q) := q$, where $\mathbf{p} \in \mathbb{Z}^d$ and $q \in \mathbb{N}$ are chosen so that $\mathbf{p}/q$ is in reduced form, it follows from Dirichlet's theorem that

$$\psi_{1+1/d} \text{ is a Dirichlet function for the triple } (\mathbb{R}^d, \mathbb{Q}^d, H), \tag{1.1}$$

where here and hereafter we write

$$\psi_c(t) := 1/t^c \tag{1.2}$$

and use the max norm to define distance on $\mathbb{R}^d$. Once a Dirichlet function has been identified, a natural question is whether it is optimal in the following sense. Call a Dirichlet function ψ *optimal* if there does not exist another Dirichlet function ϕ for which $\frac{\phi(t)}{\psi(t)} \to 0$ as $t \to \infty$, i.e., that we cannot find a faster decaying Dirichlet function. The optimality of a Dirichlet function ψ is clearly implied by[2] the set $\mathrm{BA}(M, \mathcal{Q}, H, \psi)$ of ψ*-badly approximable* points being nonempty, where we define

$$\mathrm{BA}(M, \mathcal{Q}, H, \psi) := \big\{\mathbf{x} \in M : \exists\, c_{\mathbf{x}} > 0 \text{ such that } \operatorname{dist}(\mathbf{x}, \mathbf{r}) \geq c_{\mathbf{x}} \psi\big(H(\mathbf{r})\big)\ \forall\, \mathbf{r} \in \mathcal{Q}\big\}.$$

Thus the aforementioned result of Perron shows that $\psi_{1+1/d}$ is an optimal Dirichlet function for the triple $(\mathbb{R}^d, \mathbb{Q}^d, H)$. Note that Perron's result was later strengthened by Schmidt, who showed that the set

$$\mathrm{BA}_d := \mathrm{BA}(\mathbb{R}^d, \mathbb{Q}^d, H, \psi_{1+1/d}) \tag{1.3}$$

of badly approximable vectors is of full Hausdorff dimension, and, moreover, is a winning set. As a way to interpret these results, for a Diophantine triple $(M, \mathcal{Q}, H)$ let us define the *Dirichlet exponent* $\delta(M, \mathcal{Q}, H)$ to be the supremum of $c > 0$ such that ψ_c is a Dirichlet function for $(M, \mathcal{Q}, H)$. The theorems of Dirichlet and Perron then imply that the Dirichlet exponent of $(\mathbb{R}^d, \mathbb{Q}^d, H)$ is equal to $1 + 1/d$.

Another class of examples of Diophantine triples is provided by the field of *(ambient) Diophantine approximation on manifolds* (see, for instance, [3, 4, 20]). Namely, let M be a smooth submanifold of $\mathbb{R}^d$, and consider the triple $(M, \mathbb{Q}^d, H)$. Clearly, by (1.1), $\psi_{1+1/d}$ is still a Dirichlet function. On the other hand, it is easy to choose a manifold M, for example a rational affine subspace of $\mathbb{R}^d$, such that every point of M admits a much better rate of approximation than the rate given by $\psi_{1+1/d}$. In order to rule out such behavior one is led to impose a nondegeneracy condition (see Definition 2.1). Indeed, recently, Beresnevich [4] proved that for any real-analytic nondegenerate submanifold M of $\mathbb{R}^d$, the set

$$\mathrm{BA}(M, \mathbb{Q}^d, H, \psi_{1+1/d}) = M \cap \mathrm{BA}_d$$

has full Hausdorff dimension, thereby showing the optimality of $\psi_{1+1/d}$; earlier this was established by Badziahin and Velani [2] for smooth nondegenerate planar curves. Consequently, the *ambient Dirichlet exponent*

$$\delta_M^{\mathrm{amb}} := \delta(M, \mathbb{Q}^d, H)$$

is equal to $1 + 1/d$ whenever M is nondegenerate.

[2] And in many cases equivalent to; see [13, §2] for a thorough discussion.

The goal of this paper is to develop the theory of *intrinsic approximation* on manifolds; that is, we will set (X, dist) to be $\mathbb{R}^d$ equipped with the max norm, M as a smooth submanifold of dimension $k \leq d$, $\mathcal{Q} := \mathbb{Q}^d \cap M$, and $H(\mathbf{p}/q) = q$ as before. So we are interested in the triple $(M, \mathbb{Q}^d \cap M, H)$. The field of intrinsic approximation has seen a lot of recent activity in many diverse contexts; see e.g. [6, 11, 15, 16, 23, 30]. Most recently, in the companion paper [9] we have obtained definitive results for M being a nonsingular rational quadric hypersurface of $\mathbb{R}^d$ containing a dense set of rational points. In particular, it is proved there [9, Theorem 5.1] that for such M, ψ_1 is a Dirichlet function for intrinsic approximation, and, moreover, it is optimal because [9, Theorem 4.5] when $\psi = \psi_1$, the set

$$\mathrm{BA}_M(\psi) := \mathrm{BA}(M, \mathbb{Q}^d \cap M, H, \psi)$$

has full Hausdorff dimension. Earlier this was established in [23] for the unit sphere $M = S^{d-1}$.

Now let M be an arbitrary k-dimensional nondegenerate smooth submanifold of $\mathbb{R}^d$. Is it possible to establish similar results? Clearly for that one needs some information on the set of rational points inside M. As an extreme case, Dirichlet functions do not exist if $M \cap \mathbb{Q}^d = \emptyset$. However, the following example shows that even if $\mathbb{Q}^d \cap M$ is dense in M, the "quantitative denseness" which determines Diophantine properties might depend on M.

Example 1.1. Fix $n \geq 2$, let $\Phi_n : \mathbb{R} \to \mathbb{R}^2$ be the map $x \mapsto (x, x^n)$, and let $C_n := \Phi_n(\mathbb{R})$ denote the image curve in $\mathbb{R}^2$. Then it is easy to see that $\mathbb{Q}^2 \cap C_n = \Phi_n(\mathbb{Q})$ and $H\big(\Phi_n(p/q)\big) = q^n$. Since $\psi_{1+1/d}$ is an optimal Dirichlet function on $\mathbb{R}$, this implies that $\psi_{2/n}$ is an optimal Dirichlet function for intrinsic approximation on C_n.

Note that as $n \to \infty$, the functions $\psi_{2/n}$ decay more and more slowly[3], yet none of them decays faster than ψ_1. So we can ask: does there exist a nondegenerate curve with an intrinsic Dirichlet function decaying faster than ψ_1? Our first theorem shows that the answer is no. More generally, it establishes an upper bound on the rate of decay of a Dirichlet function for intrinsic approximation on *any* k-dimensional nondegenerate submanifold $M \subseteq \mathbb{R}^d$. This is done by exhibiting for every $k \leq d$ an explicit constant $c = c(k, d)$ such that for any M as above, the set $\mathrm{BA}_M(\psi_c)$ has full Hausdorff dimension.

The constant $c(k, d)$ is arrived at via combinatorial considerations and to the authors' knowledge has not appeared previously. In some sense, it represents the heart of the paper. Given the natural way it arises in a volume computation (cf. the proof of Claim 4.2), we suspect that it will play a significant role in intrinsic Diophantine approximation moving forward. Here is how it is defined.

Notation 1.2. Denote $[n, m] := \binom{n+m}{m} = \binom{m+n}{n}$, and for $1 \leq k \leq d$, let $n_{k,d} \in \mathbb{N}$ be maximal such that

$$d = [k-1, 1] + [k-1, 2] + \ldots + [k-1, n_{k,d}] + m_{k,d} \tag{1.4}$$

for some $m_{k,d} \geq 0$, and let $m_{k,d}$ be the unique integer satisfying (1.4). Let

$$N_{k,d} := 1[k-1, 1] + 2[k-1, 2] + \ldots + n_{k,d}[k-1, n_{k,d}] + (n_{k,d} + 1)m_{k,d},$$

and define $c(k, d) := (d + 1)/N_{k,d}$.

[3]Note that the existence of these examples does not rule out the possibility that some function decaying slower than all of the functions $\psi_{2/n}$, e.g., $\psi(t) = 1/\log(t)$, is a Dirichlet function for every nondegenerate manifold whose intrinsic rationals are dense. It would be interesting to investigate this question further.

We can now state our main theorem.

Theorem 1.3. *Let $M \subseteq \mathbb{R}^d$ be a nondegenerate submanifold of dimension k. Then*

$$\dim\left(\mathrm{BA}_M(\psi_{c(k,d)})\right) = k.$$

Consequently, no Dirichlet function for intrinsic approximation on M decays faster than $\psi_{c(k,d)}$.

As a way to interpret this theorem, we can consider the *intrinsic Dirichlet exponent* of M, that is,

$$\delta_M^{\mathrm{int}} := \delta(M, \mathbb{Q}^d \cap M, H).$$

The above theorem implies that $\delta_M^{\mathrm{int}} \leq c(k,d)$ if M is a nondegenerate submanifold of $\mathbb{R}^d$ of dimension k.

The following remarks may help shed some light on the constant $c(k,d)$:

- When $k = d$, one has $n_{k,d} = 1$ and $m_{k,d} = 0$; thus $N_{k,d} = d$ and $c(k,d) = 1 + 1/d$. Therefore the $k = d$ case of Theorem 1.3 coincides with Schmidt's theorem, that is, with full Hausdorff dimension of BA_d.
- When $k = d - 1$, one has $n_{k,d} = 1$ and $m_{k,d} = 1$; thus $N_{k,d} = d + 1$ and $c(k,d) = 1$. In particular, this gives a different proof of [9, Theorem 4.5]. The latter theorem establishes full Hausdorff dimension of $\mathrm{BA}_M(\psi_1)$, and hence the optimality of Dirichlet function ψ_1, for any nonsingular rational quadric hypersurface M. Our result extends this to an arbitrary nondegenerate hypersurface M, in particular showing that whenever ψ_1 is a Dirichlet function for intrinsic approximation on M, it must be optimal.
- The reader can check that if d is fixed, then $c(k,d)$ is strictly increasing with respect to k. This confirms the intuitive logic that higher-dimensional manifolds may have more intrinsic rationals and therefore points on these manifolds should be expected to be better approximable by intrinsic rationals, so their intrinsic Dirichlet exponent should be higher. This also implies that for $k < d$, we have $c(k,d) \leq 1$. Therefore for proper nondegenerate submanifolds $M \subseteq \mathbb{R}^d$, the bound on δ_M^{int} given by Theorem 1.3 is strictly stronger than the "trivial" bound

$$\delta_M^{\mathrm{int}} \leq \delta_M^{\mathrm{amb}} = 1 + 1/d$$

  given by considering all rational points, not just intrinsic ones, and using the aforementioned result of Beresnevich on existence of badly approximable vectors on nondegenerate manifolds.
- It is also easy to compute that when $k = 1$, one has $n_{k,d} = d$ and $m_{k,d} = 0$; thus $N_{k,d} = d(d+1)/2$ and $c(1,d) = 2/d$. For other values of k, d the computation is more involved. Below is a table of values of $c(k,d)$ for $k, d \leq 6$, with rows corresponding to k and columns to d:

	1	2	3	4	5	6
1	2	1	2/3	1/2	2/5	1/3
2		3/2	1	5/6	3/4	7/11
3			4/3	1	6/7	7/9
4				5/4	1	7/8
5					6/5	1
6						7/6

Note that for some values of k, d we can construct a k-dimensional nondegenerate submanifold M of $\mathbb{R}^d$ such that $\psi_{c(k,d)}$ is a Dirichlet function for intrinsic approximation on M; for those cases, Theorem 1.3 demonstrates the optimality of this Dirichlet function and shows the constant $c(k,d)$ to be best possible. This is formalized in the following definition.

Definition 1.4. We will call a nondegenerate submanifold $M \subseteq \mathbb{R}^d$ of dimension k *maximally approximable* if $\psi_{c(k,d)}$ is a Dirichlet (and hence an optimal Dirichlet) function.

Example 1.1 shows that when $n > 2$, the curve C_n is *not* a maximally approximable submanifold of $\mathbb{R}^2$. So we know that $\psi_{c(k,d)}$ is not a Dirichlet function for some submanifolds. Nonetheless, our theorem immediately suggests the following question: for which $1 \le k \le d$ does there exist a k-dimensional maximally approximable submanifold of $\mathbb{R}^d$? We will collect examples and give some partial answers to this question in section 2.

We will prove Theorem 1.3 by showing that $\mathrm{BA}_M(\psi_{c(k,d)})$ is a winning set of a certain game. Recall that in order to prove the special case $M = \mathbb{R}^d$ of Theorem 1.3, Schmidt developed a powerful tool now known as Schmidt's game, a two-player game whose winning sets enjoy many remarkable properties, including having full dimension. In recent years, many variants of the game have been introduced, of particular note the *absolute* game of McMullen [25], and the *hyperplane* absolute game introduced in [5] (see section 3 below for more details). It is this last variant which we will utilize here. Note that a key ingredient in Schmidt's proof is the Simplex Lemma, whose original statement and proof go back to Davenport and Schmidt [29, p. 57]. A crucial step in our proof, and one which we believe to be of independent interest, is establishing an analogue of the Simplex Lemma for rationals constrained to lie on a fixed nondegenerate manifold of $\mathbb{R}^d$ (Lemma 4.1). We also develop new tools for utilizing the hyperplane absolute game, which enables us to show that $\mathrm{BA}_M(\psi_{c(k,d)})$ is hyperplane absolute winning.

Now let λ_M be a smooth volume measure on M. It is worthwhile to point out that the conclusion of Theorem 1.3, that is, full Hausdorff dimension of the set $\mathrm{BA}_M(\psi_{c(k,d)})$, cannot in general be upgraded to positive measure. Indeed, it follows from Khintchine's theorem [29, Theorem III.3A] that the Lebesgue measure of the set BA_d is zero. Also, a similar result for $\mathrm{BA}_M(\psi_1)$ where M is a nonsingular rational quadric hypersurface is a special case of [9, Theorem 6.2].

However, the situation is different if the exponent $c(k,d)$ gets replaced with a slightly bigger one. For an arbitrary Diophantine triple $(M, \mathcal{Q}, H)$ and $c > 0$ let us introduce the set

$$\begin{aligned} \mathrm{VWA}(M,\mathcal{Q},H,\psi_c) &:= M \smallsetminus \bigcup_{\varepsilon>0} \mathrm{BA}(M,\mathcal{Q},H,\psi_{c+\varepsilon}) \\ &= \left\{ \mathbf{x} \in M : \begin{array}{c} \exists\, \varepsilon > 0 \text{ and a sequence } \mathbf{r}_n \in \mathcal{Q} \text{ such that } \mathbf{r}_n \to \mathbf{x} \\ \text{and } \operatorname{dist}(\mathbf{x},\mathbf{r}_n) \le \psi_{c+\varepsilon}\big(H(\mathbf{r}_n)\big) \text{ for all } n \end{array} \right\} \end{aligned} \tag{1.5}$$

of ψ_c-*very well approximable* points. The fact that the set

$$\mathrm{VWA}_d := \mathrm{VWA}(\mathbb{R}^d, \mathbb{Q}^d, H, \psi_{1+1/d})$$

of very well approximable vectors in $\mathbb{R}^d$ is Lebesgue null is an easy consequence of the Borel–Cantelli Lemma. A similar statement for ambient approximation—namely, that $\lambda_M(\mathrm{VWA}_d) = 0$, where $M \subseteq \mathbb{R}^d$ is a nondegenerate submanifold—is much tricker. It has been conjectured by Sprindŭk in 1980 [31, Conjecture H_1] and demonstrated by Margulis and the second-named author in 1998 [20]. Later it was shown that VWA_d is μ-null for other interesting measures μ on $\mathbb{R}^d$. In particular, it follows from [18, Theorem 1.1 and Proposition 7.3] that $\mu(\mathrm{VWA}_d) = 0$ whenever μ is an *absolutely friendly* measure (see Definition 4.4) on a submanifold M as above.

Our second main theorem gives an intrinsic approximation analogue of the above statement.

Theorem 1.5 (Restated as Theorem 4.5). *Let $M \subseteq \mathbb{R}^d$ be a submanifold of dimension k. If λ_M-almost every point of M is nondegenerate, then*

$$\mathrm{VWA}_M(\psi_{c(k,d)}) := \mathrm{VWA}(M, \mathbb{Q}^d \cap M, H, \psi_{c(k,d)}) \tag{1.6}$$

is a λ_M-nullset. More generally, let $\Psi : U \to M$ be a local parameterization of M, let μ be an absolutely friendly measure on U, and let $\nu = \Psi[\mu]$. If the ν-almost every point of M is nondegenerate, then $\mathrm{VWA}_M(\psi_{c(k,d)})$ *is a ν-nullset.*

Remark 1.6. Theorem 1.5 immediately implies a result about *extrinsic* approximation, i.e., the approximation of points on a manifold M by rational points in the complement of M, described by the Diophantine triple $(M, \mathbb{Q}^d \smallsetminus M, H)$. Namely, since the exponent $c(k,d)$ is strictly less than the exponent $1 + 1/d$ appearing in Dirichlet's theorem, Theorem 1.5 implies that for almost all $\mathbf{x} \in M$, only finitely many of the approximants from Dirichlet's theorem can lie inside M, so $\mathbf{x}$ is extrinsically $\psi_{1+1/d}$-approximable. For further discussion of extrinsic approximation see [12], where a result is proven for every point in M (not just almost every point) which cannot be deduced from a corresponding "intrinsic badly approximable" result.

Question 1.7. An interesting question is whether Theorem 1.5 can be strengthened by estimating the Hausdorff dimension of $\mathrm{VWA}_M(\psi_c)$ for a fixed $c > c(k,d)$. We do this for nonsingular quadric hypersurfaces in [9] (see [10] for a generalization), where the structure of quadric hypersurfaces is explicitly used.

Outline. In section 2, we discuss some ways of constructing maximally approximable manifolds. In section 3, we recall the definition of the hyperplane absolute game and introduce two variants, which turn out to be equivalent to the original game. In section 4, we prove the main lemma of this paper, an intrinsic analogue of the Simplex Lemma, and use it to prove our main Theorems 1.3 and 1.5.

Convention 1. Throughout the paper, the symbols $\lesssim_\times$, $\gtrsim_\times$, and $\asymp_\times$ will denote multiplicative asymptotics. For example, $A \lesssim_{\times,K} B$ means that there exists a constant $C > 0$ (the *implied constant*), depending only on K, such that $A \leq CB$. In general, dependence of the implied constant on universal objects such as the manifold M will be omitted from the notation.

Convention 2. The symbol $\lhd$ will be used to indicate the end of a nested proof.

2. A discussion of maximal approximability in special cases

We start by giving a more detailed definition of a nondegenerate submanifold of $\mathbb{R}^d$.

Definition 2.1 (Cf. [20, p. 341])**.** Let $M \subseteq \mathbb{R}^d$ be a submanifold of dimension k. For each $\mathbf{x} \in M$ and $j \in \mathbb{N}$, let

$$T_{\mathbf{x}}^{(j)}(M) := \bigcup_{i=1}^{j} \{\gamma^{(i)}(0) \upharpoonright \gamma : (-\varepsilon, \varepsilon) \to M,\ \gamma(0) = \mathbf{x}\} \subseteq \mathbb{R}^d;$$

equivalently, if $\Phi : U \to M$ is a coordinate chart satisfying $\Phi(\mathbf{v}) = \mathbf{x}$,

$$T_{\mathbf{x}}^{(j)}(M) = \mathrm{Span}\left(\left\{\partial^\alpha \Phi(\mathbf{v}) : \alpha \in \mathbb{N}_0^k, 0 < |\alpha| \leq j\right\}\right). \tag{2.1}$$

Here $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$, and the power ∂^α is taken using multi-index notation: if $\alpha = (\alpha_1, \ldots, \alpha_k)$, then $\partial^\alpha = \partial_1^{\alpha_1} \cdots \partial_k^{\alpha_k}$. Also, $|\alpha| := \alpha_1 + \ldots + \alpha_k$. We will call $T_{\mathbf{x}}^{(j)}(M)$ the *tangent space of order j of M at* $\mathbf{x}$.

A point $\mathbf{x} \in M$ is said to be *D-nondegenerate* for M if $T_{\mathbf{x}}^{(D)}(M) = \mathbb{R}^d$, and *nondegenerate* if it is nondegenerate for some $D \in \mathbb{N}$. Finally, M is *nondegenerate* if some point of M is nondegenerate for M.

Observation 2.2.

(i) If M is contained in an affine hyperplane of $\mathbb{R}^d$, then M is degenerate at every point.
(ii) If M is real-analytic and connected, and M is not contained in an affine hyperplane, then every point of M is nondegenerate.
(iii) For each $D \in \mathbb{N}$, the set of D-nondegenerate points of M is relatively open in M.

Proof.

(i) If $M \subseteq \mathcal{L} + \mathbf{v}$ where $\mathcal{L}$ is a linear hyperplane and $\mathbf{v} \in \mathbb{R}^d$, then $T_{\mathbf{x}}^{(D)}(M) \subseteq \mathcal{L}$ for all $\mathbf{x} \in M$ and $D \in \mathbb{N}$.
(ii) If M is real-analytic and degenerate at $\mathbf{x} \in M$, then $M \subseteq \mathbf{x} + \bigcup_{D \in \mathbb{N}} T_{\mathbf{x}}^{(D)}(M)$.
(iii) This follows from (2.1) together with the lower semicontinuity of the function sending a matrix to its rank. $\square$

Connected manifolds which are degenerate at every point but are not contained in a hyperplane exist but are very pathological; we refer to [32] for a detailed account, stated in somewhat different language.

The next example describes an important family of nondegenerate submanifolds of $\mathbb{R}^d$.

Example 2.3 (Veronese variety)**.** Fix $k, n \in \mathbb{N}$, let $d = [k, n] - 1$ (see Notation 1.2), and consider the *Veronese embedding* $\Psi_{k,n} : \mathbb{R}^k \to \mathbb{R}^d$ defined by

$$\Psi_{k,n}(\mathbf{t}) = (\mathbf{t}^\alpha)_{\substack{\alpha \in \mathbb{N}_0^k \\ 0 < |\alpha| \leq n}},$$

where the power $\mathbf{t}^\alpha$ is taken using multi-index notation. Then it can be straightforwardly verified that the *Veronese variety* $V_{k,n} = \Psi_{k,n}(\mathbb{R}^k)$ is a nondegenerate[4] submanifold of $\mathbb{R}^d$.

[4]Even stronger, every point of $V_{k,n}$ is n-nondegenerate. Moreover, by (2.1), $\dim\left(T_{\mathbf{x}}^{(n)}(M)\right) \leq [k, n]$ for any manifold M. Thus the ambient dimension of $V_{k,n}$ is maximal among all n-nondegenerate k-dimensional manifolds.

The one-dimensional special case ($k = 1$, $d = n$) is usually called *Veronese curve* or *rational normal curve*. Letting $k = 1$ and $n = 2$ yields $V_{1,2} = C_2$ (cf. Example 1.1). Recall that the latter curve was our first example of a maximally approximable manifold (see Definition 1.4).

The following lemma shows that the map $\Psi_{k,n}$ is an "isomorphism" between the Diophantine triples $(\mathbb{R}^k, \mathbb{Q}^k, H^n)$ and $(V_{k,n}, V_{k,n} \cap \mathbb{Q}^d, H)$.

Lemma 2.4. *The map $\Psi_{k,n}$ is a diffeomorphism between $\mathbb{R}^k$ and $V_{k,n}$; moreover,*

$$V_{k,n} \cap \mathbb{Q}^{[k,n]-1} = \Psi_{k,n}(\mathbb{Q}^k) \quad \text{and} \quad H\big(\Psi_{k,n}(\mathbf{r})\big) = H^n(\mathbf{r}) \ \ \forall \mathbf{r} \in \mathbb{Q}^k.$$[5]

The proof is a straightforward computation which is left to the reader.

Corollary 2.5. *For any $k, n \in \mathbb{N}$, $V_{k,n}$ is a maximally approximable submanifold of $\mathbb{R}^d$.*

Proof. We begin by proving the following more general assertion.

Lemma 2.6. *Fix $d, n \in \mathbb{N}$, and let M be a maximally approximable submanifold of $\mathbb{R}^d$ of dimension k. Suppose that $\Psi_{d,n}(M)$ is a nondegenerate submanifold of $\mathbb{R}^{[d,n]-1}$.*[6] *Then $\Psi_{d,n}(M)$ is maximally approximable if and only if*

$$\frac{1}{n}\frac{d+1}{N_{k,d}} = \frac{[d,n]}{N_{k,[d,n]-1}}. \tag{2.2}$$

Proof. Since M is maximally approximable, $\psi_{c(k,d)}$ is an optimal Dirichlet function for the Diophantine triple $(M, M\cap\mathbb{Q}^d, H)$. Now fix $\Psi_{d,n}(\mathbf{x}) \in \Psi_{d,n}(M)$. Then there exists a sequence $M\cap\mathbb{Q}^d \ni \mathbf{r}_m \to \mathbf{x}$ with $\|\mathbf{r}_m - \mathbf{x}\| \lesssim_\times \psi_{c(k,d)} \circ H(\mathbf{r}_m)$, so by Lemma 2.4, $\Psi_{d,n}(M) \cap \mathbb{Q}^{[d,n]-1} \ni \Psi_{d,n}(\mathbf{r}_m) \to \Psi_{d,n}(\mathbf{x})$ and $\|\Psi_{d,n}(\mathbf{r}_m) - \Psi_{d,n}(\mathbf{x})\| \lesssim_\times \psi_{c(k,d)} \circ H(\mathbf{r}_m) = \psi_{c(k,d)/n} \circ H(\Psi_{d,n}(\mathbf{r}_m))$. Since $\mathbf{x}$ was arbitrary, $\psi_{c(k,d)/n}$ is a Dirichlet function for the Diophantine triple

$$\big(\Psi_{d,n}(M), \Psi_{d,n}(M) \cap \mathbb{Q}^{[d,n]-1}, H\big). \tag{2.3}$$

A similar argument shows that $\psi_{c(k,d)/n}$ is optimal for (2.3). Consequently, for any $c > 0$ the function ψ_c is an optimal Dirichlet function for (2.3) if and only if $c = c(k,d)/n$. So

$$\begin{aligned} &\Psi_{d,n}(M) \text{ is maximally approximable}\\ &\Leftrightarrow \psi_{c(k,[d,n]-1)} \text{ is an optimal Dirichlet function for (2.3)}\\ &\Leftrightarrow c(k,[d,n]-1) = c(k,d)/n\\ &\Leftrightarrow \text{(2.2) holds.} \end{aligned}$$

◁

[5]The map $\Psi_{k,n}$ is not the only embedding which is an isomorphism in this sense; more generally, if $\Psi : \mathbb{R} \to \mathbb{R}^d$ is an embedding defined by polynomials with integer coefficients, then a relation between $H \circ \Psi$ and H was found in [8, Proof of Lemma 2]. Similarly to Corollary 2.5, this relation can be used to discover an optimal Dirichlet function on the corresponding curve. However, in most cases the resulting curve is not maximally approximable.

[6]This follows, for example, if M is connected, real-analytic, and Zariski dense in $\mathbb{R}^d$.

Since $\mathbb{R}^k$ is a maximally approximable submanifold of itself, to complete the proof it suffices to show that (2.2) holds when $k = d$. Indeed,[7]

$$n_{k,k} = 1, \qquad m_{k,k} = 0, \qquad N_{k,k} = k,$$

$$n_{k,[k,n]-1} = n, \quad m_{k,[k,n]-1} = 0, \quad N_{k,[k,n]-1} = k[k+1, n-1] = \frac{kn}{k+1}[k,n],$$

which implies the desired result. □

It will be observed that Corollary 2.5 is simply the end result of transferring Dirichlet's theorem in $\mathbb{R}^k$ into $V_{k,n}$ via the map $\Psi_{k,n}$. Thus, intrinsic Diophantine approximation on $V_{k,n}$ is essentially the same as Diophantine approximation on $\mathbb{R}^k$, and does not introduce any new phenomena.

By contrast, new phenomena appear when we study intrinsic approximation on nonsingular quadric hypersurfaces, demonstrating that this theory cannot be reduced to Diophantine approximation on $\mathbb{R}^k$ in the same way. In [9] we establish a complete theory of intrinsic approximation on quadric hypersurfaces, in particular showing that ψ_1 is a Dirichlet function for every quadric hypersurface [9, Theorem 5.1], regardless of the dimension k. Since $c(k, k+1) = 1$ for every k, this shows that quadric hypersurfaces are maximally approximable.

We end this section with a discussion of the following question: For what pairs (k, d) $(1 \leq k \leq d)$ can we prove that there exists a maximally approximable submanifold of $\mathbb{R}^d$ of dimension k? For convenience let

$$\mathcal{M} = \{(k,d) : \text{ there exists a maximally approximable submanifold of } \mathbb{R}^d \text{ of dimension } k\}.$$

Trivially $(k, k) \in \mathcal{M}$ for all $k \in \mathbb{N}$. Moreover, since every nonsingular rational quadric hypersurface is maximally approximable, we have $(k, k+1) \in \mathcal{M}$ for all $k \in \mathbb{N}$. On the other hand, by Corollary 2.5, we have $(k, [k,n] - 1) \in \mathcal{M}$ for all $k, n \in \mathbb{N}$. Taking the special case $k = 1$, we have $(1, d) \in \mathcal{M}$ for all $d \in \mathbb{N}$. Thus in every dimension, there exist both a maximally approximable curve and a maximally approximable hypersurface.

It is theoretically possible to get more pairs in $\mathcal{M}$ by using Lemma 2.6. Namely, if $(k, d) \in \mathcal{M}$ and if (2.2) holds for some $n \in \mathbb{N}$, then $(k, [d, n] - 1) \in \mathcal{M}$. However, we do not have any examples of pairs (k, d) which we can prove to be in $\mathcal{M}$ this way but which were not proven to be in $\mathcal{M}$ in the above paragraph.

Although the list of pairs known to be in $\mathcal{M}$ is so far quite meager, the elegance of the calculation which produces the number $c(k, d)$ (cf. Lemma 4.1 and its proof) leads the authors to believe that there could be many more examples. It is even conceivable that all dimension pairs are in $\mathcal{M}$.

The smallest pair (k, d) for which we do not know the answer to this question is the pair $(2, 4)$, which satisfies $c(2, 4) = 5/6$.

3. The hyperplane game and two variants

In [28], W. M. Schmidt introduced the game which is now known as Schmidt's game. A variant of this game was defined by C. T. McMullen [25], and in turn a variant of McMullen's game was defined in [5]. For the purposes of this paper, we

[7]In establishing these formulas, the identity $[a, b] = [a - 1, b] + [a, b - 1]$ is useful.

will be interested only in this last variant, called the *hyperplane absolute game*,[8] and not in Schmidt's game or McMullen's game. However, we note that every hyperplane winning set is winning for Schmidt's game [5, Proposition 2.3(a)]. Some recent papers in which the hyperplane game has appeared are [1, 19, 26].

Given $\beta > 0$ and $k \in \mathbb{N}$, the β-hyperplane game is played on $\mathbb{R}^k$ by two players, Alice and Bob, as follows:

1. Bob chooses an initial ball $B_0 = B(\mathbf{x}_0, \rho_0) \subseteq \mathbb{R}^k$. (In this paper all balls are closed.)
2. After Bob's nth move B_n, Alice chooses an affine hyperplane $A_n \subseteq \mathbb{R}^k$. We say that Alice "deletes the neighborhood of A_n".
3. After Alice's nth move A_n, Bob chooses a ball $B_{n+1} = B(\mathbf{x}_{n+1}, \rho_{n+1})$ satisfying
$$B_{n+1} \subseteq B_n \setminus A_n^{(\beta\rho_n)} \text{ and } \rho_{n+1} \geq \beta\rho_n.$$
Here and elsewhere $S^{(\varepsilon)}$ denotes the closed ε-thickening of a set S. If he is unable to choose such a ball, he loses.[9]

A set $S \subseteq \mathbb{R}^k$ is said to be *β-hyperplane winning* if Alice has a strategy which guarantees that
$$\bigcap_{n=1}^{\infty} B_n \cap S \neq \emptyset.$$
Note that if $\beta < \beta'$, then any β-hyperplane winning set is β'-hyperplane winning. We say that S is *hyperplane winning* if it is β-hyperplane winning for all $\beta > 0$, or equivalently, if there exist arbitrarily small $\beta > 0$ such that S is β-hyperplane winning.

Remark 3.1. By modifying slightly the proof of [14, Proposition 4.4], one can show that if Bob's balls are required to satisfy $\rho_{n+1} = \beta\rho_n$ rather than $\rho_{n+1} \geq \beta\rho_n$, then the class of sets which are hyperplane winning remains unchanged. Thus we can assume that $\rho_n \to 0$, in which case the intersection $\bigcap_1^\infty B_n$ is a singleton.

We list here three important results regarding hyperplane winning sets, the proofs of which can be found in [5, Proposition 2.3(b,c), Lemma 4.1, and Proposition 4.7].

Proposition 3.2.

(i) *The countable intersection of hyperplane winning sets is hyperplane winning.*
(ii) *The image of a hyperplane winning set under a $\mathcal{C}^1$ diffeomorphism of $\mathbb{R}^k$ is hyperplane winning.*
(iii) *The intersection of any hyperplane winning set with any open set has full Hausdorff dimension.*

[8] In what follows we abbreviate "hyperplane absolute game" to just "hyperplane game".

[9] This disagrees with the convention introduced in [5]; however, if we restrict to $0 < \beta \leq 1/3$ (as is done in [5]), then Bob is always able to make a legal move, so the question is irrelevant. We use the convention that Bob loses in order to avoid technicalities (cf. [7, p. 4]) in the variants of the hyperplane game discussed below, where it is not always obvious whether or not Bob has legal moves. Our convention allows us to ignore the issue in the sense that we do not need to check whether Bob has legal moves whenever we are trying to prove that a certain strategy of Alice's is winning.

In fact, in (iii) more is true: the intersection of a hyperplane winning set with a sufficiently nondegenerate fractal (a *hyperplane diffuse* set) is winning for Schmidt's game on that fractal [5, Propositions 4.7 and 4.9] and therefore has Hausdorff dimension equal to at least the lower pointwise dimension of any measure whose support is equal to that fractal [21, Proposition 5.1]. In particular, if the fractal is Ahlfors regular then the intersection has full dimension relative to the fractal.

In [22, §3], the notion of hyperplane winning was generalized from subsets of Euclidean space to subsets of arbitrary manifolds. Namely, a subset S of a manifold M is *hyperplane winning relative to* M if whenever $\Psi : U \to M$ is a local parameterization of M and $K \subseteq U$ is compact, the set $\Psi^{-1}(S) \cup (\mathbb{R}^k \setminus K)$ is hyperplane winning.[10]

We now state our main result concerning the abundance of badly intrinsically approximable points.

Theorem 3.3 (Restated as Theorem 4.3)**.** *Let $M \subseteq \mathbb{R}^d$ be a submanifold of dimension k, and let $c(k,d)$ be as in Notation* 1.2. *Suppose that for some $D \in \mathbb{N}$, every point of M is D-nondegenerate. Then* $\mathrm{BA}_M(\psi_{c(k,d)})$ *is hyperplane winning relative to M.*

Using Theorem 3.3, we deduce as a corollary a theorem stated in the introduction:

Proof of Theorem 1.3 *using Theorem* 3.3. For each $D \in \mathbb{N}$, let $M_D \subseteq M$ be the set of D-nondegenerate points of M. Since M is nondegenerate, there exists $D \in \mathbb{N}$ such that $M_D \neq \emptyset$; then $\mathrm{BA}_M(\psi_{c(k,d)}) \cap M_D$ is hyperplane winning relative to M_D. By (iii) of Proposition 3.2, $\mathrm{BA}_M(\psi_{c(k,d)}) \cap M_D$ has Hausdorff dimension k. □

In order to prove Theorem 3.3, we will introduce two variants of the hyperplane game. The first allows Alice to delete neighborhoods of algebraic sets rather than just hyperplanes, and the second allows her to delete neighborhoods of levelsets of smooth functions. It will turn out that each of these variants is equivalent to the hyperplane game, meaning that any set which is winning for one of the games is winning for all three games.

Definition 3.4. Fix $\beta > 0$ and $D \in \mathbb{N}$. The rules of the (β, D) *algebraic-set game* are the same as the rules of the β-hyperplane game, except that A_n is allowed to be the zero set of any nonzero polynomial of degree at most D. A set is *algebraic-set winning* if there exists $D \in \mathbb{N}$ so that it is (β, D) algebraic-set winning for all $\beta > 0$.

Given a ball $B \subseteq \mathbb{R}^k$ and a $\mathcal{C}^D$ function $f : B \to \mathbb{R}$, for each $\mathbf{x} \in B$ let

$$\|f\|_{\mathcal{C}^D,\mathbf{x}} := \max_{\substack{\alpha \in \mathbb{N}_0^k \\ |\alpha| \leq D}} |f^{(\alpha)}(\mathbf{x})| \,,$$

where the derivative is taken using multi-index notation. Let

$$\|f\|_{\mathcal{C}^D,B} := \sup_{\mathbf{x} \in B} \|f\|_{\mathcal{C}^D,\mathbf{x}} \,.$$

[10] In [22] the definition is given in a slightly different way, depending on the notion of hyperplane winning subsets of an open set. However, it is easily verified that the two definitions are equivalent.

Definition 3.5. The rules of the (β, D, C_1)-*levelset game* are the same as the rules of the β-hyperplane game, except that A_n is allowed to be the zero set of any nonzero $\mathcal{C}^{D+1}$ function $f : B_n \to \mathbb{R}$ satisfying

$$\|f\|_{\mathcal{C}^{D+1},B_n} \leq C_1 \|f\|_{\mathcal{C}^D,B_n}. \tag{3.1}$$

A set is *levelset winning* if there exist $D \in \mathbb{N}$ and $C_1 > 0$ so that it is (β, D, C_1)-levelset winning for all $\beta > 0$.

The condition (3.1) should be interpreted heuristically as meaning that "f is close to being a polynomial of degree D".

Clearly, any hyperplane winning set is algebraic-set winning and any algebraic-set winning set is levelset winning. The remainder of this section is devoted to the proof of the following theorem.

Theorem 3.6. *Any levelset winning set is hyperplane winning.*

We begin by introducing some notation.

Notation 3.7.

- For $f : U \to \mathbb{R}$, Z_f will denote the zero set of f, i.e. $Z_f = f^{-1}(0)$.
- For $D \in \mathbb{N}$, $\mathcal{P}_D$ will denote the set of all polynomials of degree at most D whose largest coefficient has magnitude 1. Note that $\mathcal{P}_D$ is a compact topological space; moreover, every nonzero polynomial of degree at most D is a nonzero scalar multiple of an element of $\mathcal{P}_D$.

Lemma 3.8. *Fix $k \in \mathbb{N}$ and $0 < \beta \leq 1$, and let $f : \mathbb{R}^k \to \mathbb{R}$ be a nonzero polynomial. Suppose that Bob and Alice are playing the β-hyperplane game, and suppose that Bob's first move is $B_0 = B(\mathbf{0}, 1)$. Then there exists $\gamma > 0$ so that Alice has a strategy to guarantee that Bob's first ball of radius less than γ (assuming that such a ball exists) is disjoint from $Z_f^{(\gamma)}$.*

Proof. The proof is by induction on the degree of f. If $\deg(f) = 0$, then f is a nonzero constant and $Z_f = \emptyset$, so the lemma is trivially satisfied. Next, suppose that the lemma is true for all polynomials of degree strictly less than the degree of f. In particular, it is true for $\widetilde{f} := \partial_i f$, where $i = 1, \ldots, d$ is chosen so that $\widetilde{f}$ is nonzero. Let $\widetilde{\gamma} > 0$ be the value given by the induction hypothesis. Since $L := Z_f \setminus Z_{\widetilde{f}}$ is a smooth $(k-1)$-dimensional submanifold of $\mathbb{R}^k$, for all $\mathbf{x} \in L$ and for all sufficiently small $\rho > 0$, $L \cap B(\mathbf{x}, \rho)$ is contained in $\mathcal{L}^{(\beta\rho/4)}$ for some hyperplane $\mathcal{L} \subseteq \mathbb{R}^k$ (specifically, the tangent plane of L at $\mathbf{x}$). Thus since $K := Z_f \cap B(\mathbf{0}, 2) \setminus \mathrm{Int}(Z_{\widetilde{f}}^{(\widetilde{\gamma}/2)})$ is a compact subset of L, there exists $\delta > 0$ with the following property:

$$\begin{aligned}&\text{For every ball } B(\mathbf{x}, \rho) \subseteq \mathbb{R}^k \text{ satisfying } 0 < \rho \leq \delta,\\ &\text{there exists a hyperplane } \mathcal{L} \subseteq \mathbb{R}^k \text{ such that } K \cap B(\mathbf{x}, 2\rho) \subseteq \mathcal{L}^{(\beta\rho/2)}.\end{aligned} \tag{3.2}$$

(If $K \cap B(\mathbf{x}, 2\rho) \neq \emptyset$, then $\mathcal{L}$ may be chosen to be the tangent plane of L at some point of $K \cap B(\mathbf{x}, 2\rho)$.) Let $\gamma = \beta^2 \min(\widetilde{\gamma}, \delta)/2$.

Alice's strategy is now as follows: Use the strategy from the induction hypothesis to guarantee that Bob's first ball of radius less than $\widetilde{\gamma}$ is disjoint from $Z_{\widetilde{f}}^{(\widetilde{\gamma})}$. If the radius of this ball is greater than δ, make further moves arbitrarily until Bob chooses a ball of radius less than δ. Either way, let $B = B(\mathbf{x}, \rho)$ denote Bob's first ball satisfying $\rho \leq \min(\widetilde{\gamma}, \delta)$, and note that $\rho \geq \beta \min(\widetilde{\gamma}, \delta) = 2\gamma/\beta$. In particular $\rho > \gamma$, so Bob has not yet chosen a ball of radius less than γ. Let $\mathcal{L}$ be a hyperplane such that $K \cap B(\mathbf{x}, 2\rho) \subseteq \mathcal{L}^{(\beta\rho/2)}$, guaranteed to exist by (3.2). Alice's next move will be to delete the $\beta\rho$-neighborhood of the hyperplane $\mathcal{L}$. Following that, she will make arbitrary moves until Bob chooses a ball $\widetilde{B}$ of radius less than γ.

We claim that $\widetilde{B}$ is disjoint from $Z_f^{(\gamma)}$. Indeed, fix $\mathbf{y} \in \widetilde{B} \subseteq B \setminus \mathcal{L}^{(\beta\rho)}$. Then for $\mathbf{z} \in Z_f$, either

(1) $\mathbf{z} \in K \cap B(\mathbf{x}, 2\rho) \subseteq \mathcal{L}^{(\beta\rho/2)}$, in which case

$$\|\mathbf{z} - \mathbf{y}\| \geq \mathrm{dist}(\mathcal{L}^{(\beta\rho/2)}, \mathbb{R}^k \setminus \mathcal{L}^{(\beta\rho)}) = \beta\rho/2 \geq \gamma,$$

(2) $\mathbf{z} \notin B(\mathbf{x}, 2\rho)$, in which case

$$\|\mathbf{z} - \mathbf{y}\| \geq \mathrm{dist}\left(\mathbb{R}^k \setminus B(\mathbf{x}, 2\rho), B(\mathbf{x}, \rho)\right) = \rho \geq \gamma,$$

(3) $\mathbf{z} \notin B(\mathbf{0}, 2)$, in which case

$$\|\mathbf{z} - \mathbf{y}\| \geq \mathrm{dist}\left(\mathbb{R}^k \setminus B(\mathbf{0}, 2), B(\mathbf{0}, 1)\right) = 1 \geq \gamma, \text{ or}$$

(4) $\mathbf{z} \in Z_f \cap B(\mathbf{0}, 2) \setminus K \subseteq Z_{\widetilde{f}}^{(\widetilde{\gamma}/2)}$, in which case

$$\|\mathbf{z} - \mathbf{y}\| \geq \mathrm{dist}(Z_{\widetilde{f}}^{(\widetilde{\gamma}/2)}, \mathbb{R}^k \setminus Z_{\widetilde{f}}^{(\widetilde{\gamma})}) \geq \widetilde{\gamma}/2 \geq \gamma.$$

Thus $\mathbf{y} \notin Z_f^{(\gamma)}$. □

We next show that the constant γ can be made to depend only on the degree of f and not on f itself.

Lemma 3.9. *Fix $k, D \in \mathbb{N}$ and $0 < \beta \leq 1$. There exists $\gamma > 0$ such that for any nonzero polynomial $f : \mathbb{R}^k \to \mathbb{R}$ of degree at most D, if Bob and Alice play the β-hyperplane game and if Bob's first move is $B_0 = B(\mathbf{0}, 1)$, then Alice has a strategy to guarantee that Bob's first ball of radius less than γ (assuming that such a ball exists) is disjoint from $Z_f^{(\gamma)}$.*

Proof. The map $\mathcal{P}_D \ni f \mapsto Z_f$ is upper semicontinuous in the Vietoris topology (cf. [17, §4.F]), meaning that for any $f \in \mathcal{P}_D$, $\gamma > 0$, and $K \subseteq \mathbb{R}^k$ compact, there exists a neighborhood of f in $\mathcal{P}_D$ such that all g in the neighborhood satisfy $Z_g \cap K \subseteq Z_f^{(\gamma)}$. In particular, for each $f \in \mathcal{P}_D$, let γ_f be as in Lemma 3.8, and let $U_f \subseteq \mathcal{P}_D$ be a neighborhood of f such that for all $g \in U_f$, $Z_g \cap B(\mathbf{0}, 2) \subseteq Z_f^{(\gamma_f/2)}$. Since $\mathcal{P}_D$ is compact, there exists a finite sequence $(f_i)_{i=1}^n$ such that the collection $(U_{f_i})_{i=1}^n$ covers $\mathcal{P}_D$. Let $\gamma = \min_{i=1}^n \gamma_{f_i}/2$. Then for all $g \in \mathcal{P}_D$, we have $g \in U_{f_i}$ for some i, and so

$$Z_g^{(\gamma)} \cap B(\mathbf{0}, 1) \subseteq Z_{f_i}^{(\gamma_{f_i}/2+\gamma)} \subseteq Z_{f_i}^{(\gamma_{f_i})}.$$

Since Alice has a strategy to avoid $Z_{f_i}^{(\gamma_{f_i})}$ by the time Bob's radius is less than γ_{f_i}, she has a strategy to avoid $Z_g^{(\gamma)}$ by the time Bob's radius is less than γ. □

Let k, D, β, and γ be as above. Fix $\mathbf{x} \in \mathbb{R}^k$ and $\rho > 0$, and let

$$T_{\mathbf{x},\rho}(\mathbf{w}) = \mathbf{x} + \rho\mathbf{w}, \tag{3.3}$$

so that $T_{\mathbf{x},\rho}\big(B(\mathbf{0}, 1)\big) = B(\mathbf{x}, \rho)$. Translating Lemma 3.8 via the map $T_{\mathbf{x},\rho}$, we see that if Bob and Alice play the β-hyperplane game, then after Bob makes a move $B(\mathbf{x}, \rho)$, Alice may devote the next several turns to ensuring that Bob's first ball of radius less than $\gamma\rho$ is disjoint from $Z_{f \circ T_{\mathbf{x},\rho}^{-1}}^{(\gamma\rho)}$. This allows her to translate any winning strategy for the $(\beta\gamma, D)$ algebraic-set game into a winning strategy for the

β-hyperplane game. Indeed, if in the algebraic-set game Alice responds to Bob's move $B(\mathbf{x},\rho)$ by deleting the set $Z_f^{(\gamma\rho)}$, then in the β-hyperplane game Alice simply spends the next several turns avoiding $Z_f^{(\gamma\rho)}$. Bob's first ball of radius less than $\gamma\rho$ will still have radius $\geq \beta\gamma\rho$ by the rules of the β-hyperplane game, so it can be interpreted as Bob's next move in the $(\beta\gamma, D)$ algebraic set game. Summarizing, we have the following corollary.

Corollary 3.10. *Any algebraic-set winning set is hyperplane winning.*

Proof. For each $k, D \in \mathbb{N}$ and $0 < \beta \leq 1$, if $\gamma > 0$ is as in Lemma 3.9, then every $(\beta\gamma, D)$ algebraic-set winning subset of $\mathbb{R}^k$ is β-hyperplane winning. $\square$

To complete the proof of Theorem 3.6, we must show that every levelset winning set is algebraic-set winning. For this, we will need three more lemmas.

Lemma 3.11. *Fix $k, D \in \mathbb{N}$ and $\beta > 0$. Then there exists $\gamma > 0$ such that for any $f \in \mathcal{P}_D$, there exists $g \in \mathcal{P}_D$ such that*

$$f^{-1}(-\gamma,\gamma) \cap B(\mathbf{0},1) \subseteq Z_g^{(\beta)}.$$

Proof. For each $g \in \mathcal{P}_D$, $|g|$ is bounded uniformly away from 0 on $B(\mathbf{0},1) \setminus Z_g^{(\beta)}$. Let $\gamma_g > 0$ be strictly less than this uniform bound, and let U_g be the set of all polynomials $f \in \mathcal{P}_D$ such that $\min_{B(\mathbf{0},1)\setminus Z_g^{(\beta)}} |f| > \gamma_g$. Then U_g is an open set containing g. Letting $(U_{g_i})_{i=1}^n$ be a finite subcover, the lemma holds with $\gamma = \min_{i=1}^n \gamma_{g_i}$. $\square$

Lemma 3.12. *Fix $k, D \in \mathbb{N}$ and $\beta > 0$, and let $B = B(\mathbf{0},1)$. There exists $\delta > 0$ such that if $f : B \to \mathbb{R}$ is a $\mathcal{C}^{D+1}$ function satisfying*

$$\sup_B |f^{(\alpha)}| \leq \delta \|f\|_{\mathcal{C}^D,B} \;\; \forall \alpha \in \mathbb{N}_0^k \text{ with } |\alpha| = D+1,$$

then there exists $g \in \mathcal{P}_D$ such that

$$Z_f \subseteq f^{-1}(-\delta\|f\|_{\mathcal{C}^D,B}, \delta\|f\|_{\mathcal{C}^D,B}) \subseteq Z_g^{(\beta)}.$$

In particular, $Z_f^{(\beta)} \subseteq Z_g^{(2\beta)}$.

Proof. Fix $\delta > 0$ small to be determined, and let $f : B \to \mathbb{R}$ be as above. For convenience of notation, without loss of generality we assume that $\|f\|_{\mathcal{C}^D,B} = 1$. By the definition of $\|f\|_{\mathcal{C}^D,B}$, there exists a point $\mathbf{z} \in B$ such that $\|f\|_{\mathcal{C}^D,\mathbf{z}} \geq 1/2$. Let $h_{\mathbf{z}}$ denote the Dth order Taylor polynomial for f centered at $\mathbf{z}$. Then

$$\|h_{\mathbf{z}}\|_{\mathcal{C}^D,B} \geq \|h_{\mathbf{z}}\|_{\mathcal{C}^D,\mathbf{z}} = \|f\|_{\mathcal{C}^D,\mathbf{z}} \geq 1/2\,. \tag{3.4}$$

By Taylor's theorem, for all $\mathbf{x} \in B$,

$$|f(\mathbf{x}) - h_{\mathbf{z}}(\mathbf{x})| \lesssim_\times \max_{|\alpha|=D+1} \sup_B |f^{(\alpha)}| \cdot \|\mathbf{x}-\mathbf{z}\|^{D+1} \lesssim_\times \delta\,,$$

and so

$$|h_{\mathbf{z}}(\mathbf{x})| \lesssim_\times \delta \;\; \forall \mathbf{x} \in f^{-1}(-\delta,\delta)\,.$$

Write $h_{\mathbf{z}} = cj$ for some $c > 0$ and $j \in \mathcal{P}_D$; then $\|j\|_{\mathcal{C}^D,B} \asymp_\times 1$ since $\mathcal{P}_D$ is compact. Combining with (3.4), we see that $c \gtrsim_\times 1$, and thus

$$|j(\mathbf{x})| \lesssim_\times \delta \;\; \forall \mathbf{x} \in f^{-1}(-\delta,\delta). \tag{3.5}$$

Let $\gamma > 0$ be as in Lemma 3.11, and let δ be γ divided by the implied constant of (3.5). Then

$$f^{-1}(-\delta, \delta) \subseteq j^{-1}(-\gamma, \gamma).$$

Moreover, by Lemma 3.11 there exists $g \in \mathcal{P}_D$ such that $j^{-1}(-\gamma, \gamma) \subseteq Z_g^{(\beta)}$. This completes the proof. □

Lemma 3.13. *Fix $k, D \in \mathbb{N}$ and $\beta, C_1 > 0$. Then there exists $\varepsilon > 0$ such that for any ball $B = B(\mathbf{x}, \rho) \subseteq \mathbb{R}^k$ satisfying $\rho \leq \varepsilon$ and for any $\mathcal{C}^{D+1}$ function $f : B \to \mathbb{R}$ satisfying*

$$\|f\|_{\mathcal{C}^{D+1}, B} \leq C_1 \|f\|_{\mathcal{C}^D, B},$$

there exists a polynomial $g : \mathbb{R}^k \to \mathbb{R}$ of degree at most D such that $Z_f \subseteq Z_g^{(\beta\rho)}$, and thus $Z_f^{(\beta\rho)} \subseteq Z_g^{(2\beta\rho)}$.

Proof. Fix $0 < \varepsilon \leq 1$ small to be determined, and let $B = B(\mathbf{x}, \rho)$ and $f : B \to \mathbb{R}$ be as above. Let $T_{\mathbf{x},\rho}$ be given by (3.3), and let $\widetilde{f} = f \circ T_{\mathbf{x},\rho}$. Then for all $\alpha \in \mathbb{N}_0^k$ with $|\alpha| = D + 1$,

$$\sup_{B(\mathbf{0},1)} |\widetilde{f}^{(\alpha)}| = \rho^{D+1} \sup_{B(\mathbf{0},1)} |f^{(\alpha)} \circ T_{\mathbf{x},\rho}| \leq \rho^{D+1} \|f\|_{\mathcal{C}^{D+1}, B} \lesssim_\times \rho^{D+1} \|f\|_{\mathcal{C}^D, B},$$

and on the other hand

$$\|\widetilde{f}\|_{\mathcal{C}^D, B(\mathbf{0},1)} = \max_{|\alpha| \leq D} \sup_{B(\mathbf{0},1)} \widetilde{f}^{(\alpha)} = \max_{|\alpha| \leq D} \rho^{|\alpha|} \sup_{B(\mathbf{0},1)} |f^{(\alpha)} \circ T_{\mathbf{x},\rho}| \geq \rho^D \|f\|_{\mathcal{C}^D, B}.$$

Combining, we have

$$\sup_{B(\mathbf{0},1)} |\widetilde{f}^{(\alpha)}| \lesssim_\times \rho \|\widetilde{f}\|_{\mathcal{C}^D, B(\mathbf{0},1)} \;\; \forall \alpha \in \mathbb{N}_0^k \text{ with } |\alpha| = D + 1.$$

So for ε sufficiently small, $\widetilde{f}$ satisfies the hypotheses of Lemma 3.12. Let $\widetilde{g}$ be the polynomial given by Lemma 3.12, and let $g = \widetilde{g} \circ T_{\mathbf{x},\rho}^{-1}$, so that $Z_g = T_{\mathbf{x},\rho}(Z_{\widetilde{g}})$. This completes the proof. □

Let k, D, β, C_1, and ε be as above. Lemma 3.13 gives us a way of translating a winning strategy for Alice in the (β, D, C_1)-levelset game into a winning strategy for Alice in the $(2\beta, D)$ algebraic-set game. Indeed, without loss of generality suppose that Bob's first move in the (β, D, C_1)-levelset game has radius $\leq \varepsilon$. (Otherwise Alice makes dummy moves until this is true.) Now if Alice responds to Bob's move $B(\mathbf{x}, \rho)$ in the (β, D, C_1)-levelset game by deleting the set $Z_f^{(\beta\rho)}$, then in the $(2\beta, D)$ algebraic-set game, she will simply delete the set $Z_g^{(2\beta\rho)}$, where g is given by Lemma 3.13. Summarizing, we have the following corollary.

Corollary 3.14. *Any levelset winning set is algebraic-set winning.*

Proof. For each $k, D \in \mathbb{N}$ and $\beta, C_1 > 0$, then every (β, D, C_1)-levelset winning subset of $\mathbb{R}^k$ is $(2\beta, D)$ algebraic-set winning. □

Combining Corollaries 3.10 and 3.14 completes the proof of Theorem 3.6.

4. The simplex lemma and its consequences

The paradigmatic example of a hyperplane winning set is the set BA_d defined by the formula (1.3), which was proven to be hyperplane winning in [5, Theorem 2.5], as a consequence of the so-called *simplex lemma* [5, Lemma 3.1]. Essentially, the simplex lemma states that for each ball $B(\mathbf{x}, \rho) \subseteq \mathbb{R}^d$, the set of rational points in $B(\mathbf{x}, \rho)$ whose denominators are less than $\varepsilon\rho^{-d/(d+1)}$ is contained in an affine hyperplane, where $\varepsilon > 0$ is small and depends only on d. As a result, when playing the hyperplane game Alice can simply delete the neighborhood of the hyperplane given by the simplex lemma, and it turns out that this strategy is winning for BA_d. In this section we prove an analogue of the simplex lemma for rational points in a fixed manifold M. We then use the simplex lemma to prove two general negative results about intrinsic approximation on manifolds: that $\mathrm{BA}_M(\psi_{c(k,d)})$ is hyperplane winning, and that $\lambda_M(\mathrm{VWA}_M) = 0$.

Lemma 4.1 (Simplex lemma for intrinsic approximation on manifolds)**.** *Let $M \subseteq \mathbb{R}^d$ be a submanifold of dimension k, let $\Psi : U \to M$ be a local parameterization of M, and let $V \subseteq U$ be compact. Then there exists $\kappa > 0$ such that for all $\mathbf{s} \in U$ and $0 < \rho \leq 1$, the set*

$$S_{\mathbf{s},\rho} := \{\mathbf{p}/q \in \mathbb{Q}^d \cap \Psi\big(V \cap B(\mathbf{s}, \rho)\big) : q \leq \kappa\rho^{-1/c(k,d)}\}$$

is contained in a hyperplane.

Proof. For all $\mathbf{s} \in U$ let $\Phi(\mathbf{s}) = (1, \Psi(\mathbf{s}))$. Define a function $f : U^{d+1} \to \mathbb{R}$ by

$$f(\mathbf{s}_1, \ldots, \mathbf{s}_{d+1}) = \det[\Phi(\mathbf{s}_1) \ \cdots \ \Phi(\mathbf{s}_{d+1})].$$

Then f vanishes along the diagonal

$$\Delta = \{(\mathbf{t}, \ldots, \mathbf{t}) : \mathbf{t} \in U\}.$$

In fact, the first several derivatives of f vanish along the diagonal, due to repeated columns:

Claim 4.2. *The smallest order derivative of f which does not vanish identically along the diagonal is no less than $N_{k,d}$.*

Proof. Suppose that for some sequence of multi-indices $\alpha_1, \ldots, \alpha_{d+1} \in \mathbb{N}_0^k$, the expression

$$\frac{\partial}{\partial^{\alpha_1}\mathbf{t}_1} \cdots \frac{\partial}{\partial^{\alpha_{d+1}}\mathbf{t}_{d+1}} f(\mathbf{t}_1, \ldots, \mathbf{t}_{d+1}) \tag{4.1}$$

does not vanish identically along the diagonal $\{\mathbf{t}_1 = \cdots = \mathbf{t}_{d+1}\}$. Here we use the multi-index notation

$$\frac{\partial}{\partial^{\alpha_i}\mathbf{t}_i} = \left(\frac{\partial}{\partial t_{i,1}}\right)^{\alpha_{i,1}} \cdots \left(\frac{\partial}{\partial t_{i,k}}\right)^{\alpha_{i,k}}.$$

Since the determinant of a matrix is linear with respect to the columns of that matrix, we have

$$\frac{\partial}{\partial^{\alpha_1}\mathbf{t}_1} \cdots \frac{\partial}{\partial^{\alpha_{d+1}}\mathbf{t}_{d+1}} f(\mathbf{t}_1, \ldots, \mathbf{t}_{d+1}) = \det[\partial^{\alpha_1}\Phi(\mathbf{t}_1) \ \cdots \ \partial^{\alpha_{d+1}}\Phi(\mathbf{t}_{d+1})].$$

Since this does not vanish identically along the diagonal, there exists $\mathbf{t} \in U$ such that

$$\det[\partial^{\alpha_1}\Phi(\mathbf{t}) \ \cdots \ \partial^{\alpha_{d+1}}\Phi(\mathbf{t})] \neq 0.$$

In particular, the rows $(\partial^{\alpha_i}\Phi(\mathbf{t}))_{i=1}^{d+1}$ are all distinct, so the multi-indices $\alpha_1,\ldots,\alpha_{d+1}$ must be distinct. Thus for each $j\in\mathbb{N}$,

$$n_j := \#\{i=1,\ldots,d+1 : |\alpha_i| = j\} \leq \#\{\alpha\in\mathbb{N}_0^k : |\alpha| = j\} = [k-1,j], \tag{4.2}$$

and, on the other hand,

$$\sum_{j=0}^{\infty} n_j = d+1. \tag{4.3}$$

The order of the derivative (4.1) is

$$\sum_{i=1}^{d+1}|\alpha_i| = \sum_{j=1}^{\infty} j n_j,$$

so computing the smallest order derivative of f which potentially does not vanish along the diagonal becomes a combinatorial problem of minimizing $\sum_{j=1}^{\infty} j n_j$ subject to (4.2) and (4.3). The reader will verify that the minimum is attained when

$$n_j = \begin{cases} [k-1,j] & \text{if } j < n_{k,d}+1,\\ m_{k,d} & \text{if } j = n_{k,d}+1 \qquad (j\geq 0),\\ 0 & \text{if } j > n_{k,d}+1,\end{cases}$$

and that the value of $\sum_{j=1}^{\infty} j n_j$ on this sequence is $N_{k,d}$, where $n_{k,d}$, $m_{k,d}$, and $N_{k,d}$ are as in Notation 1.2. $\triangleleft$

Thus by Taylor's theorem and the compactness of V, we have

$$|f(\mathbf{s}_1,\ldots,\mathbf{s}_{d+1})| \lesssim_\times \operatorname{dist}\left((\mathbf{s}_i)_1^{d+1},\Delta\right)^{N_{k,d}} \tag{4.4}$$

for all $\mathbf{s}_1,\ldots,\mathbf{s}_{d+1}\in V$.

Now by contradiction, suppose that the points $\mathbf{r}_1,\ldots,\mathbf{r}_{d+1}\in S_{\mathbf{s},\rho}$ do not lie in a hyperplane. For each i write $\mathbf{r}_i = \Psi(\mathbf{s}_i) = \mathbf{p}_i/q_i$. Let

$$D = f(\mathbf{s}_1,\ldots,\mathbf{s}_{d+1}) = \det\begin{bmatrix} 1 & \cdots & 1\\ \mathbf{r}_1 & \cdots & \mathbf{r}_{d+1}\end{bmatrix} \neq 0.$$

Since $\mathbf{s}_i\in B(\mathbf{s},\rho)$, we have $\operatorname{dist}\left((\mathbf{s}_i)_1^{d+1},\Delta\right)\lesssim_\times \rho$. Thus by (4.4) we have

$$|D| \lesssim_\times \rho^{N_{k,d}}. \tag{4.5}$$

On the other hand, we have

$$D = \prod_{i=1}^{d+1}\frac{1}{q_i}\det\begin{bmatrix} q_1 & \cdots & q_{d+1}\\ \mathbf{p}_1 & \cdots & \mathbf{p}_{d+1}\end{bmatrix} \in \prod_{i=1}^{d+1}\frac{1}{q_i}\mathbb{Z}.$$

Thus,

$$|D| \geq \frac{1}{\prod_{i=1}^{d+1} q_i}.$$

Since by assumption $q_i\leq \kappa\rho^{-N_{k,d}/(d+1)}$, we have

$$|D| \geq \kappa^{-(d+1)}\rho^{N_{k,d}}.$$

For $\kappa>0$ sufficiently small, this contradicts (4.5). $\square$

Using the simplex lemma, we proceed to prove two results about intrinsic approximation on M. The first is the following.

Theorem 4.3. *Let $M \subseteq \mathbb{R}^d$ be a submanifold of dimension k, and let $c(k,d)$ be as in Notation 1.2. Suppose that for some $D \in \mathbb{N}$, every point of M is D-nondegenerate. Then $\mathrm{BA}_M(\psi_{c(k,d)})$ is hyperplane winning relative to M.*

Proof. Let $\Psi : U \to M$ be a local parameterization of M, and let $K \subseteq U$ be compact. We need to show that the following set is hyperplane winning:

$$\Psi^{-1}\big(\mathrm{BA}_M(\psi_{c(k,d)})\big) \cup (\mathbb{R}^k \setminus K). \tag{4.6}$$

Fix $C_1 > 0$ large to be determined, and let $\beta > 0$. We will show that the set (4.6) is (β, D, C_1)-levelset winning, where D is as in the statement of Theorem 4.3. Let $\lambda = \beta^{-1/c(k,d)}$ (so that $\lambda > 1$). Denote Bob's first move by $B_0 = B(\mathbf{s}_0, \rho_0) \subseteq \mathbb{R}^k$. Fix an open set $V \supseteq K$ which is relatively compact in U; without loss of generality we may assume that $B(\mathbf{s}_0, 2\rho_0) \subseteq V$, since Alice may make dummy moves until either this is true or Bob's ball is disjoint from K. Now Alice's strategy is as follows: If Bob has just made his nth move $B_n = B(\mathbf{s}_n, \rho_n) \subseteq V$, then Alice will delete the $\beta\rho_n$-neighborhood of the set $\Psi^{-1}(\mathcal{L}_n)$, where $\mathcal{L}_n$ is the affine hyperplane containing the set $S_{\mathbf{s}_n, 2\rho_n}$. To complete the proof we need to show (i) that this is legal (given $C_1 > 0$ large enough), and (ii) that the strategy guarantees that $\bigcap_1^\infty B_n \cap \Psi^{-1}\big(\mathrm{BA}_M(\psi_{c(k,d)})\big) \neq \emptyset$.

(i) For each $\mathbf{s} \in U$ let $\Phi(\mathbf{s}) = (1, \Psi(\mathbf{s}))$. Since every point of M is D-nondegenerate, for each $\mathbf{s} \in U$ and $\mathbf{w} \in \mathbb{R}^{d+1} \setminus \{\mathbf{0}\}$ we have

$$\|\mathbf{t} \mapsto \mathbf{w} \cdot \Phi(\mathbf{t})\|_{\mathcal{C}^D, \mathbf{s}} > 0,$$

and by continuity, this quantity is bounded from below uniformly for $\mathbf{s} \in V$ and $\mathbf{w} \in S^d$.

Now consider Alice's nth move. Write $\mathcal{L}_n = \{\mathbf{x} \in \mathbb{R}^d : \mathbf{w} \cdot (1, \mathbf{x}) = 0\}$ for some $\mathbf{w} \in S^d$. Define $f : U \to \mathbb{R}$ by

$$f(\mathbf{s}) = \mathbf{w} \cdot \Phi(\mathbf{s}),$$

so that $Z_f = \Psi^{-1}(\mathcal{L}_n)$. Then by the first paragraph, $\|f\|_{\mathcal{C}^D, B_n} \geq \|f\|_{\mathcal{C}^D, \mathbf{s}_n}$ is bounded from below. On the other hand,

$$\|f\|_{\mathcal{C}^{D+1}, B_n} \leq \|\mathbf{w}\| \cdot \|\Phi\|_{\mathcal{C}^{D+1}, B_n} \leq \|\Phi\|_{\mathcal{C}^{D+1}, V} \asymp_\times 1,$$

so $\|f\|_{\mathcal{C}^{D+1}, B_n} \lesssim_\times \|f\|_{\mathcal{C}^D, B_n}$. Letting C_1 be the implied constant finishes the proof.

(ii) By Remark 3.1, we may assume that $\bigcap_1^\infty B_n$ is a singleton, say $\bigcap_1^\infty B_n = \{\mathbf{s}\}$. For each $\mathbf{r} = \mathbf{p}/q \in \mathbb{Q}^d \cap M$, let $n \in \mathbb{N}$ be minimal such that $q \leq \kappa\rho_n^{-N_{k,d}/(d+1)}$. If $\mathbf{r} \in \Psi\big(V \cap B(\mathbf{s}_n, 2\rho_n)\big)$, then by Lemma 4.1 we have $\mathbf{r} \in \mathcal{L}_n$. Since Alice deleted the set $\Psi^{-1}(\mathcal{L}_n)^{(\beta\rho_n)}$, we have $\mathbf{s} \notin \Psi^{-1}(\mathcal{L}_n)^{(\beta\rho_n)}$ and thus

$$\mathrm{dist}(\Psi(\mathbf{s}), \mathbf{r}) \asymp_\times \|\mathbf{s} - \Psi^{-1}(\mathbf{r})\| \geq \beta\rho_n \asymp_{\times,\beta,\kappa} q^{-(d+1)/N_{k,d}}.$$

On the other hand, if $\mathbf{r} \notin \Psi\big(V \cap B(\mathbf{s}_n, 2\rho_n)\big)$, then either $\mathbf{r} \notin \Psi(V)$, which implies

$$\mathrm{dist}(\Psi(\mathbf{s}), \mathbf{r}) \geq \mathrm{dist}\,\big(\Psi(\mathbf{s}), M \setminus \Psi(V)\big) \asymp_\times 1,$$

or $\mathbf{r} \in \Psi\big(V \setminus B(\mathbf{s}_n, 2\rho_n)\big)$, in which case

$$\mathrm{dist}(\Psi(\mathbf{s}), \mathbf{r}) \asymp_\times \|\mathbf{s} - \Psi^{-1}(\mathbf{r})\| \geq \rho_n \asymp_{\times,\beta,\kappa} q^{-(d+1)/N_{k,d}}.$$

In all cases we have $\mathrm{dist}(\Psi(\mathbf{s}), \mathbf{r}) \gtrsim_\times \psi_{c(k,d)}(q)$, so $\Psi(\mathbf{s}) \in \mathrm{BA}_M(\psi_{c(k,d)})$. $\square$

To state our last theorem regarding general manifolds, we need a definition.

Definition 4.4. A measure μ on an open set $U \subseteq \mathbb{R}^k$ is *absolutely decaying* if there exist $C, \alpha > 0$ such that for all $\mathbf{x} \in \mathrm{Supp}(\mu)$, for all $0 < \rho \leq 1$ such that $B(\mathbf{x}, \rho) \subseteq U$, for all $\varepsilon > 0$, and for every affine hyperplane $\mathcal{L} \subseteq \mathbb{R}^k$, we have

$$\mu\left(\mathcal{L}^{(\varepsilon\rho)} \cap B(\mathbf{x}, \rho)\right) \leq C\varepsilon^\alpha \mu\big(B(\mathbf{x}, \rho)\big).$$

We call μ *doubling* if $\mu\big(B(\mathbf{x}, 2\rho)\big) \asymp_\times \mu\big(B(\mathbf{x}, \rho)\big)$ for all $\mathbf{x} \in \mathrm{Supp}(\mu)$ and $0 < \rho \leq 1$. If μ is both absolutely decaying and doubling, then μ is called *absolutely friendly.*

Theorem 4.5. *Let $M \subseteq \mathbb{R}^d$ be a submanifold of dimension k, and fix $D \in \mathbb{N}$. Let $\Psi : U \to M$ be a local parameterization of M, let μ be an absolutely friendly measure on U, and let $\nu = \Psi[\mu]$. If ν-a.e. point of M is D-nondegenerate, then* $\mathrm{VWA}_M(\psi_{c(k,d)})$ *is a ν-nullset. In particular, $\lambda_M\big(\mathrm{VWA}_M(\psi_{c(k,d)})\big) = 0$.*

Proof. Let $\kappa > 0$ be as in Lemma 4.1. Fix $\lambda > 1$ arbitrary, and for each $n \in \mathbb{N}$ let

$$T_n := \kappa\lambda^{n/c(k,d)},$$
$$\rho_n := \frac{1}{2}\lambda^{-n}.$$

Let $K \subseteq U$ be a compact set, and let $V \supseteq K$ be open and relatively compact in U. Then by Lemma 4.1, we have the following corollary.

Corollary 4.6. *For all $\mathbf{s} \in V$ and for all $n \in \mathbb{N}$, the set*

$$S_{n,\mathbf{s}} = \big\{\mathbf{p}/q \in \mathbb{Q}^d \cap \Psi\big(V \cap B(\mathbf{s}, 2\rho_n)\big) : q \leq T_n\big\}$$

is contained in a hyperplane.

Denote the hyperplane guaranteed by Corollary 4.6 by $\mathcal{L}_{n,\mathbf{s}}$. For each $n \in \mathbb{N}$, let $(\mathbf{s}_i^{(n)})_{i=1}^{N_n}$ be a maximal ρ_n-separated subset of K. Then $\{B(\mathbf{s}_i^{(n)}, \rho_n) : i = 1, \ldots, N_n\}$ is a cover of K whose multiplicity is bounded depending only on d. For each $i = 1, \ldots, N_n$, let $\mathcal{L}_{n,i} = \mathcal{L}_{n,\mathbf{s}_i^{(n)}}$.

Claim 4.7.

$$\Psi^{-1}\big(\mathrm{VWA}_M(\psi_{c(k,d)})\big) \cap K \subseteq \lim_{\gamma\to 0} \limsup_{n\to\infty} \bigcup_{i=1}^{N_n} \left[\big(\Psi^{-1}(\mathcal{L}_{n,i})\big)^{(\rho_n^{1+\gamma})} \cap B(\mathbf{s}_i^{(n)}, \rho_n)\right].$$

Proof. Fix $\mathbf{s} \in \Psi^{-1}\big(\mathrm{VWA}_M(\psi_{c(k,d)})\big) \cap K$, and recall the definition of the set $\mathrm{VWA}_M(\psi_{c(k,d)})$ given by (1.5) and (1.6). Since $\Psi(\mathbf{s}) \in \mathrm{VWA}_M(\psi_{c(k,d)})$, there exists $\varepsilon > 0$ such that there are infinitely many $\mathbf{r} = \mathbf{p}/q \in \mathbb{Q}^d \cap M$ satisfying

$$\|\Psi(\mathbf{s}) - \mathbf{r}\| \leq q^{-(c(k,d)+\varepsilon)}. \tag{4.7}$$

Fix such an $\mathbf{r}$, and let $n \in \mathbb{N}$ satisfy $T_{n-1} \leq q < T_n$. Then

$$\|\mathbf{s} - \Psi^{-1}(\mathbf{r})\| \asymp_\times \|\Psi(\mathbf{s}) - \mathbf{r}\| \leq T_{n-1}^{-(c(k,d)+\varepsilon)} \asymp_\times \rho_n^{1+\varepsilon/c(k,d)}.$$

Fix $0 < \gamma < \varepsilon/c(k,d)$. If n is sufficiently large, then we have

$$\|\mathbf{s} - \Psi^{-1}(\mathbf{r})\| \leq \rho_n^{1+\gamma} \leq \rho_n.$$

On the other hand, since $\mathbf{s} \in K$, we have $\mathbf{s} \in B(\mathbf{s}_i^{(n)}, \rho_n)$ for some $i = 1, \ldots, N_n$. It follows that $\mathbf{r} \in \Psi\big(V \cap B(\mathbf{s}_i^{(n)}, 2\rho_n)\big)$, and so by Corollary 4.6 we have $\mathbf{r} \in \mathcal{L}_{n,i}$. Thus

$$\mathbf{s} \in \big(\Psi^{-1}(\mathcal{L}_{n,i})\big)^{(\rho_n^{1+\gamma})} \cap B(\mathbf{s}_i^{(n)}, \rho_n).$$

Since this argument holds for all $\mathbf{r}$ satisfying (4.7), it follows that

$$\mathbf{s} \in \bigcup_{i=1}^{N_n} \left[\big(\Psi^{-1}(\mathcal{L}_{n,i})\big)^{(\rho_n^{1+\gamma})} \cap B(\mathbf{s}_i^{(n)}, \rho_n)\right]$$

for infinitely many $n \in \mathbb{N}$. $\triangleleft$

Claim 4.8. *For each $\gamma > 0$, there exists $\alpha > 0$ such that for all $n \in \mathbb{N}$ and $i = 1, \ldots, N_n$,*

$$\mu\left(\Psi^{-1}(\mathcal{L}_{n,i})^{(\rho_n^{1+\gamma})} \cap B(\mathbf{s}_i^{(n)}, \rho_n)\right) \lesssim_\times \rho_n^\alpha \mu\big(B(\mathbf{s}_i^{(n)}, \rho_n)\big). \tag{4.8}$$

Proof. For each $\mathbf{s} \in U$ let $\Phi(\mathbf{s}) = (1, \Psi(\mathbf{s}))$. Let $B = B(\mathbf{s}_i^{(n)}, \rho_n)$. By [18, Proposition 7.3], there exist $C, \alpha > 0$ such that for any linear map $P : \mathbb{R}^{d+1} \to \mathbb{R}$, if $f = P \circ \Phi$, then

$$\mu\left(f^{-1}\left(-\rho_n^{\gamma/2} \sup_B |f|, \rho_n^{\gamma/2} \sup_B |f|\right) \cap B\right) \leq \rho_n^\alpha \mu(B). \tag{4.9}$$

On the other hand, if P is the linear functional whose zero set is the hyperplane $\mathcal{L}_{n,i}$, then

$$\Psi^{-1}(\mathcal{L}_{n,i}) = Z_f.$$

So to complete the proof, we must show that

$$Z_f^{(\rho_n^{1+\gamma})} \cap B \subseteq f^{-1}\left(-\rho_n^{\gamma/2} \sup_B |f|, \rho_n^{\gamma/2} \sup_B |f|\right). \tag{4.10}$$

Let $T = T_{\mathbf{s}_i^{(n)}, \rho_n}$ be as in (3.3), and let $\widetilde{f} = f \circ T$. Translating (4.10) via T gives

$$Z_{\widetilde{f}}^{(\rho_n^\gamma)} \cap B(\mathbf{0}, 1) \subseteq \widetilde{f}^{-1}\left(-\rho_n^{\gamma/2} \sup_{B(\mathbf{0},1)} |\widetilde{f}|, \rho_n^{\gamma/2} \sup_{B(\mathbf{0},1)} |\widetilde{f}|\right). \tag{4.11}$$

To demonstrate (4.11), we observe that if $\mathbf{s} \in Z_{\widetilde{f}}^{(\rho_n^\gamma)} \cap B(\mathbf{0}, 1)$, then there exists $\mathbf{t} \in Z_{\widetilde{f}}$ for which $\|\mathbf{s} - \mathbf{t}\| \leq \rho_n^\gamma$. By Taylor's theorem, we have

$$|\widetilde{f}(\mathbf{s})| \lesssim_\times \|\widetilde{f}\|_{\mathcal{C}^1, B(\mathbf{0},2)} \|\mathbf{s} - \mathbf{t}\| \leq \|\widetilde{f}\|_{\mathcal{C}^D, B(\mathbf{0},2)} \rho_n^\gamma.$$

So to complete the proof, we must show that

$$\|\widetilde{f}\|_{\mathcal{C}^D, B(\mathbf{0},2)} \lesssim_\times \sup_{B(\mathbf{0},1)} |\widetilde{f}|. \tag{4.12}$$

To demonstrate (4.12), let $\beta > 0$ be small enough so that for every polynomial g of degree at most D, $B(\mathbf{0},1) \not\subseteq Z_g^{(2\beta)}$. Such a β exists, e.g., by a compactness argument. Let $\delta > 0$ be given by Lemma 3.12. For n sufficiently large, the argument of Lemma 3.13 shows that the hypotheses of Lemma 3.12 are satisfied for $\widetilde{f}$, and thus that

$$\widetilde{f}^{-1}(-\delta\|\widetilde{f}\|_{\mathcal{C}^D,B(\mathbf{0},2)}, \delta\|\widetilde{f}\|_{\mathcal{C}^D,B(\mathbf{0},2)}) \subseteq Z_g^{(2\beta)} \subsetneqq B(\mathbf{0},1).$$

Thus there exists $\mathbf{s} \in B(\mathbf{0},1)$ for which $|\widetilde{f}(\mathbf{s})| \geq \delta\|\widetilde{f}\|_{\mathcal{C}^D,B(\mathbf{0},2)}$, demonstrating (4.12). $\triangleleft$

Fix γ, α as in Claim 4.8. From (4.8), we see that

$$\sum_{i=1}^{N_n} \mu\left(\left(\Psi^{-1}(\mathcal{L}_{n,i})\right)^{(\rho_n^{1+\gamma})} \cap B(\mathbf{s}_i^{(n)}, \rho_n)\right) \lesssim_\times \rho_n^\alpha \sum_{i=1}^{N_n} \mu\big(B(\mathbf{s}_i^{(n)}, \rho_n)\big) \lesssim_\times \rho_n^\alpha \mu(V),$$

$$\sum_{n=0}^{\infty}\sum_{i=1}^{N_n} \mu\left(\left(\Psi^{-1}(\mathcal{L}_{n,i})\right)^{(\rho_n^{1+\gamma})} \cap B(\mathbf{s}_i^{(n)}, \rho_n)\right) \lesssim_\times \sum_{n=0}^{\infty} \lambda^{-\alpha n} < \infty.$$

Thus by the Borel–Cantelli lemma, for each $\gamma > 0$,

$$\mu\left(\limsup_{n\to\infty} \bigcup_{i=1}^{N_n} \left[\left(\Psi^{-1}(\mathcal{L}_{n,i})\right)^{(\rho_n^{1+\gamma})} \cap B(\mathbf{s}_i^{(n)}, \rho_n)\right]\right) = 0,$$

and so $\mu\big(\mathrm{VWA}_M(\psi_{c(k,d)})\big) = 0$ by Claim 4.7. $\square$

Acknowledgements

The authors would like to thank Victor Beresnevich, Yann Bugeaud, and Sanju Velani for helpful discussions, and the anonymous referee for valuable comments.

References

[1] Jinpeng An, Lifan Guan, and Dmitry Kleinbock, *Bounded orbits of diagonalizable flows on $SL_3(R)/SL_3(Z)$*, Int. Math. Res. Not. IMRN **24** (2015), 13623–13652, DOI 10.1093/imrn/rnv120. MR3436158

[2] Dzmitry Badziahin and Sanju Velani, *Badly approximable points on planar curves and a problem of Davenport*, Math. Ann. **359** (2014), no. 3-4, 969–1023, DOI 10.1007/s00208-014-1020-z. MR3231023

[3] Victor Beresnevich, *Rational points near manifolds and metric Diophantine approximation*, Ann. of Math. (2) **175** (2012), no. 1, 187–235, DOI 10.4007/annals.2012.175.1.5. MR2874641

[4] Victor Beresnevich, *Badly approximable points on manifolds*, Invent. Math. **202** (2015), no. 3, 1199–1240, DOI 10.1007/s00222-015-0586-8. MR3425389

[5] Ryan Broderick, Lior Fishman, Dmitry Kleinbock, Asaf Reich, and Barak Weiss, *The set of badly approximable vectors is strongly C^1 incompressible*, Math. Proc. Cambridge Philos. Soc. **153** (2012), no. 2, 319–339, DOI 10.1017/S0305004112000242. MR2981929

[6] Ryan Broderick, Lior Fishman, and Asaf Reich, *Intrinsic approximation on Cantor-like sets, a problem of Mahler*, Mosc. J. Comb. Number Theory **1** (2011), no. 4, 3–12. MR2988380

[7] Ryan Broderick, Lior Fishman, and David Simmons, *Badly approximable systems of affine forms and incompressibility on fractals*, J. Number Theory **133** (2013), no. 7, 2186–2205, DOI 10.1016/j.jnt.2012.12.004. MR3035957

[8] Natalia Budarina, Detta Dickinson, and Jason Levesley, *Simultaneous Diophantine approximation on polynomial curves*, Mathematika **56** (2010), no. 1, 77–85, DOI 10.1112/S0025579309000382. MR2604984

[9] Lior Fishman, Dmitry Kleinbock, Keith Merrill, and David Simmons, *Intrinsic Diophantine approximation on quadric hypersurfaces*, `http://arxiv.org/abs/1405.7650`, preprint 2015.

[10] Lior Fishman, Keith Merrill, and David Simmons, *Hausdorff dimensions of very well intrinsically approximable subsets of quadratic hypersurfaces*, `http://arxiv.org/abs/1502.07648`, preprint 2015.

[11] Lior Fishman and David Simmons, *Intrinsic approximation for fractals defined by rational iterated function systems: Mahler's research suggestion*, Proc. Lond. Math. Soc. (3) **109** (2014), no. 1, 189–212, DOI 10.1112/plms/pdu002. MR3237740

[12] Lior Fishman and David Simmons, *Extrinsic Diophantine approximation on manifolds and fractals*, J. Math. Pures Appl. (9) **104** (2015), no. 1, 83–101, DOI 10.1016/j.matpur.2015.02.002. MR3350721

[13] Lior Fishman, David Simmons, and Mariusz Urbański, *Diophantine approximation in Banach spaces*, J. Théor. Nombres Bordeaux **26** (2014), no. 2, 363–384. MR3320484

[14] Lior Fishman, David Simmons, and Mariusz Urbański, *Diophantine approximation and the geometry of limit sets in Gromov hyperbolic metric spaces*, `http://arxiv.org/abs/1301.5630`, preprint 2013, to appear Mem. Amer. Math. Soc.

[15] Anish Ghosh, Alexander Gorodnik, and Amos Nevo, *Diophantine approximation and automorphic spectrum*, Int. Math. Res. Not. IMRN **21** (2013), 5002–5058. MR3123673

[16] Anish Ghosh, Alexander Gorodnik, and Amos Nevo, *Metric Diophantine approximation on homogeneous varieties*, Compos. Math. **150** (2014), no. 8, 1435–1456, DOI 10.1112/S0010437X13007859. MR3252026

[17] Alexander S. Kechris, *Classical descriptive set theory*, Graduate Texts in Mathematics, vol. 156, Springer-Verlag, New York, 1995. MR1321597

[18] Dmitry Kleinbock, Elon Lindenstrauss, and Barak Weiss, *On fractal measures and Diophantine approximation*, Selecta Math. (N.S.) **10** (2004), no. 4, 479–523, DOI 10.1007/s00029-004-0378-2. MR2134453

[19] Dmitry Kleinbock and Tue Ly, *Badly approximable S-numbers and absolute Schmidt games*, J. Number Theory **164** (2016), 13–42, DOI 10.1016/j.jnt.2015.12.014. MR3474376

[20] D. Y. Kleinbock and G. A. Margulis, *Flows on homogeneous spaces and Diophantine approximation on manifolds*, Ann. of Math. (2) **148** (1998), no. 1, 339–360, DOI 10.2307/120997. MR1652916

[21] Dmitry Kleinbock and Barak Weiss, *Modified Schmidt games and Diophantine approximation with weights*, Adv. Math. **223** (2010), no. 4, 1276–1298, DOI 10.1016/j.aim.2009.09.018. MR2581371

[22] Dmitry Kleinbock and Barak Weiss, *Values of binary quadratic forms at integer points and Schmidt games*, Recent trends in ergodic theory and dynamical systems, Contemp. Math., vol. 631, Amer. Math. Soc., Providence, RI, 2015, pp. 77–92, DOI 10.1090/conm/631/12597. MR3330339

[23] Dmitry Kleinbock and Keith Merrill, *Rational approximation on spheres*, Israel J. Math. **209** (2015), no. 1, 293–322, DOI 10.1007/s11856-015-1219-z. MR3430242

[24] Adrien-Marie Legendre, *Essai sur la théorie des nombres* (French), Cambridge Library Collection, Cambridge University Press, Cambridge, 2009. Reprint of the second (1808) edition. MR2859036

[25] Curtis T. McMullen, *Winning sets, quasiconformal maps and Diophantine approximation*, Geom. Funct. Anal. **20** (2010), no. 3, 726–740, DOI 10.1007/s00039-010-0078-3. MR2720230

[26] Erez Nesharim and David Simmons, $\mathbf{Bad}(s,t)$ *is hyperplane absolute winning*, Acta Arith. **164** (2014), no. 2, 145–152, DOI 10.4064/aa164-2-4. MR3224831

[27] Oskar Perron, *Über diophantische Approximationen* (German), Math. Ann. **83** (1921), no. 1-2, 77–84, DOI 10.1007/BF01464229. MR1512000

[28] Wolfgang M. Schmidt, *On badly approximable numbers and certain games*, Trans. Amer. Math. Soc. **123** (1966), 178–199, DOI 10.2307/1994619. MR0195595

[29] Wolfgang M. Schmidt, *Diophantine approximation*, Lecture Notes in Mathematics, vol. 785, Springer, Berlin, 1980. MR568710

[30] Eric Schmutz, *Rational points on the unit sphere*, Cent. Eur. J. Math. **6** (2008), no. 3, 482–487, DOI 10.2478/s11533-008-0038-4. MR2425007

[31] V. G. Sprindžuk, *Achievements and problems of the theory of Diophantine approximations* (Russian), Uspekhi Mat. Nauk **35** (1980), no. 4(214), 3–68, 248. MR586190

[32] K. Wolsson, *Linear dependence of a function set of m variables with vanishing generalized Wronskians*, Linear Algebra Appl. **117** (1989), 73–80, DOI 10.1016/0024-3795(89)90548-X. MR993032

Department of Mathematics, University of North Texas, 1155 Union Circle #311430, Denton, Texas 76203-5017

E-mail address: lior.fishman@unt.edu

Department of Mathematics, Brandeis University, 415 South Street, Waltham, Massachusetts 02454-9110

E-mail address: kleinboc@brandeis.edu

Department of Mathematics, Brandeis University, 415 South Street, Waltham, Massachusetts 02454-9110

E-mail address: merrill2@brandeis.edu

Department of Mathematics, University of York, Heslington, York YO10 5DD, United Kingdom

E-mail address: David.Simmons@york.ac.uk