SOME REMARKS ON MAHLER'S CLASSIFICATION IN HIGHER DIMENSION

S. KRISTENSEN, S. H. PEDERSEN, B. WEISS

ABSTRACT. We prove a number of results on the metric and non-metric theory of Diophantine approximation for Yu's multidimensional variant of Mahler's classification of transcendental numbers.

1. INTRODUCTION

In [11], Mahler introduced a classification of transcendental numbers in terms of their approximation properties by algebraic numbers. More precisely, he introduced for each $k \in \mathbb{N}$ and each $\alpha \in \mathbb{R}$ the Diophantine exponent

$$\omega_k(x) = \sup\{\omega \in \mathbb{R} : |P(x)| \leq H(P)^{-\omega} \text{ for infinitely many irreducible } P \in \mathbb{Z}[X], \deg(P) \leq k\}. \quad (1)$$

Here, $H(P)$ denotes the naive height of the polynomial P, i.e. the maximum absolute value among the coefficients of P.

Mahler defined classes of numbers according to the asymptotic behaviour of these exponents as k increases. More precisely, let

$$\omega(x) = \limsup_{k\to\infty} \frac{\omega_k(x)}{k}.$$

The number x belongs to one of the following four classes.

- x is an A-number if $\omega(x) = 0$, so that x is algebraic over $\mathbb{Q}$.
- x is an S-number if $0 < \omega(x) < \infty$.
- x is a T-number if $\omega(x) = \infty$, but $\omega_k(x) < \infty$ for all k.
- x is a U-number if $\omega(x) = \infty$ and $\omega_k(x) = \infty$ for all k large enough.

All four classes are non-empty, with almost all real numbers being S-numbers. Every real number belongs to one of the classes, and the classes are invariant under algebraic operations over $\mathbb{Q}$.

In analogy with Mahler's classification, Koksma [10] introduced a different classification based on the exponent

$$\omega_k^*(\alpha) = \sup\{\omega^* \in \mathbb{R} : |x - \alpha| \leq H(\alpha)^{-\omega^*} \text{ for infinitly many } \alpha \in \overline{\mathbb{Q}} \cap \mathbb{R}, \deg(\alpha) \leq k\}.$$

In this case, $H(\alpha)$ denotes the naive height of α, i.e. the naive height of the minimal integer polynomial of α. In analogy with Mahler's classification, one defines $w^*(x)$ and A^*-, S^*-, T^*- and U^*-numbers.

The reader is referred to the monograph [4] for an excellent overview of the classifications and their properties. A particular property is that the classifications coincide, so that A-numbers are A^*-numbers, S-numbers are S^*-numbers and so on. The individual exponents however need not coincide.

2010 *Mathematics Subject Classification.* 11J82, 11J83.

SK and SHP supported by the Danish Natual Science Research Council.

In [16], Yu introduced a classification similar to Mahler's for d-tuples of real numbers. In brief, the classification is completely similar, except that the exponents $\omega_k(x)$ are now defined in terms of integer polynomials in d variables.

An analogue of Koksma's classification was introduced by Schmidt [14]. However, the relation between the two classifications is not at all clear, and it is conjectured that the two classifications do not agree [14].

It is the purpose of the present note to study the Diophantine approximation problems arising within Yu's classification. We recall the simple connection between the questions arising from Mahler's classification, and the problem of diophantine approximation with dependent quantities. A classical problem in Diophantine approximation, given $\mathbf{x} = (x_1, \dots, x_d) \in \mathbb{R}^d$, is to to find ω for which

$$\|\mathbf{q} \cdot \mathbf{x}\| \le (\max_{1 \le i \le d} |q_i|)^{-\omega} \text{ for infinitely many } \mathbf{q} = (q_1, \dots, q_d) \in \mathbb{Z}^d, \tag{2}$$

where as usual $\| \cdot \|$ denotes the distance to the nearest integer. Comparing (1) and (2), one sees that one can define Mahler's exponents ω_k by restricting the classical problem to a consideration of vectors $\mathbf{x}$ belonging to the Veronese curve

$$\Gamma = \left\{ (x^1, \dots, x^k) \in \mathbb{R}^k : x \in \mathbb{R} \right\}.$$

Similarly, in order to understand the exponents arising in Yu's classification, one should once more consider the corresponding problem of a single linear form, but replace the Veronese curve by the variety obtained by letting the coordinates consist of the distinct monomials in d variables of degree at most k, say. The resulting Diophantine approximation properties considered in this case would correspond to the multidimensional analogue of ω_k, i.e.

$$\omega_k(\mathbf{x}) = \sup\{\omega \in \mathbb{R} : |P(\mathbf{x})| \le H(P)^{-\omega} \text{ for infinitely many } P \in \mathbb{Z}[X_1, \dots, X_d], \deg(P) \le k\}.$$

Throughout, let $n = \binom{k+d}{d} - 1$ be the number of nonconstant monomials in d variables of total degree at most k. In addition to the usual, naive height $H(P)$, we will also use the following modification $\tilde{H}(P)$, which is the maximum absolute value of the coefficients of the non-contant terms of P. The following is a slight re-statement of [16, Theorem 1].

Theorem 1. *For any $\mathbf{x} = (x_1, \dots, x_d) \in \mathbb{R}^d$, there exists $c(k, \mathbf{x}) > 0$ such that for all $Q > 1$, there is a polynomial $P \in \mathbb{Z}[X_1, \dots, X_d]$ of total degree at most k and height $H(P) \le Q$, such that*

$$|P(\mathbf{x})| < c(k, \mathbf{x}) Q^{-n}.$$

Replacing the condition $H(P) \le Q$ by $\tilde{H}(P) \le Q$, we may always choose $c(k, \mathbf{x}) = 1$.

The proof is essentially an application of the pigeon hole principle, and is completely analogous to the classical proof of Dirichlet's approximation theorem in higher dimension. As a standard corollary, one obtains the first bounds on the exponents $\omega_k(\mathbf{x})$.

Corollary 2. *For any $\mathbf{x} = (x_1, \dots, x_d) \in \mathbb{R}^d$, there exists a $c(k, \mathbf{x}) > 0$ such that*

$$|P(\mathbf{x})| < c(k, \mathbf{x}) H(P)^{-n},$$

for infinitely many $P \in \mathbb{Z}[X_1, \dots, X_d]$ of total degree at most k. In particular, $\omega_k(\mathbf{x}) \ge n$.

The corollary tells us what the normalising factor in the multidimensional definition of $\omega(x)$ should be, namely the number of non-constant monomials in d variables of total degree at most k.

Inspired by the above result, we will define the notions of k-very well approximable, k-badly approximable, k-singular and k-Dirichlet improvable. We will then proceed to prove that the set defined in this manner are all Lebesgue null-sets and so are indeed exceptional. In the case of k-badly approximable results, we will also show that these form a thick set, i.e. a set whose intersection with any ball has maximal Hausdorff dimension. In fact, many of our results are somewhat stronger than these statements. The properties are all consequences of other work by various authors (see below). Finally, we will deduce a Roth type theorem from Schmidt's Subspace Theorem [13].

2. Results and proofs

In each of the following subsections we introduce a property of approximation of d-tuples of real numbers by alebraic numbers, and prove a result about it which extends previous results known in case $d = 1$.

2.1. **k-very well approximable points.** A point $\mathbf{x} = (x_1, \dots, x_d) \in \mathbb{R}^d$ is called k-very well approximable if there exists $\varepsilon > 0$ and infinitely many polynomials $P \in \mathbb{Z}[X_1, \dots, X_d]$ of total degree at most k, such that

$$|P(\mathbf{x})| \leq H(P)^{-(n+\varepsilon)}. \tag{3}$$

In other words, $\mathbf{x}$ is k-very well approximable if the exponent n on the right hand side in Corollary 2 can be increased by a positive amount. We will prove that this property is exceptional in the sense that almost no points with respect to the d-dimensional Lebesgue measure are k-very well approximable. In fact, we will show that this property is stable under restriction to subsets supporting a measure with nice properties.

We recall some properties of measures from [7]. A measure μ on $\mathbb{R}^d$ is said to be *Federer* (or doubling) if there is a number $D > 0$ such that for any $x \in \operatorname{supp}(\mu)$ and any $r > 0$, the ball $B(x, r)$ centered at x of radius r satisfies

$$\mu\big(B(x, 2r)\big) < D\mu\big(B(x, r)\big). \tag{4}$$

The measure μ is said to be *absolutely decaying* if for some pair of numbers $C, \alpha > 0$

$$\mu\left(B(x, r) \cap \mathcal{L}^{(\varepsilon)}\right) \leq C \left(\frac{\varepsilon}{r}\right)^{\alpha} \mu\big(B(x, r)\big), \tag{5}$$

for any ball $B(x, r)$ with $x \in \operatorname{supp}(\mu)$ and any affine hyperplane $\mathcal{L}$, where $\mathcal{L}^{(\varepsilon)}$ denotes the ε-neighbourhood of $\mathcal{L}$. A weaker variant of the property of being absolutely decaying is obtained by replacing r in the denominator on the right hand side of (5) by the quantity

$$\sup\{c > 0 : \mu(\{z \in B(x, r) : \operatorname{dist}(z, \mathcal{L}) > c\}) > 0\}.$$

In this case, we say that μ is *decaying.* If the measure μ has the property that

$$\mu(\mathcal{L}) = 0, \tag{6}$$

for any affine hyperplane $\mathcal{L}$, μ is called *non-planar.* Note that an absolutely decaying measure is automatically non-planar, but a decaying measure need not be non-planar. Finally, μ is called *absolutely friendly* if it Federer and absolutely decaying, and is called *friendly* if it is Federer, decaying, and non-planar.

Theorem 3. *Let μ be an absolutely decaying Federer measure on $\mathbb{R}^d$. For any $k \in \mathbb{N}$, the set of k-very well approximable points is a null set with respect to μ. In particular, Lebesgue almost-no points are k-very well approximable.*

Our proof relies on results of [7], in which the case $d = 1$ was proved.

Proof. Let $f : \mathbb{R}^d \to \mathbb{R}^n$ be defined by $f(x_1, \dots, x_d) = (x_1, x_2, \dots, x_{d-1}x_d^{k-1}, x_d^k)$, so that f maps $(x_1, \dots, x_d)$ to the n distinct nonconstant monomials in d variables of total degree at most k. Clearly, f is smooth, and by taking partial derivatives, we easily see that $\mathbb{R}^n$ may be spanned by the partial derivatives of f of order up to k.

From [7, Theorem 2.1(b)] we immediately see that the pushforward $f_*\mu$ is a friendly measure on $\mathbb{R}^n$. We now apply [7, Theorem 1.1], which states that a friendly measure is strongly extremal, i.e. for any $\delta > 0$, almost no points in the support of the measure have the property that

$$\prod_{i=1}^{n} |qy_i - p_i| < q^{-(1+\delta)},$$

for infinitely many $\mathbf{p} \in \mathbb{Z}^n$, $q \in \mathbb{N}$. Clearly, this implies the weaker property of extremality, i.e. that for any $\delta' > 0$, almost no points in the support of the measure satisfy

$$\max_{1 \le i \le n} |qy_i - p_i| < q^{-(\frac{1}{n}+\delta')}, \tag{7}$$

for infinitely many $\mathbf{p} \in \mathbb{Z}^n$, $q \in \mathbb{N}$.

To get from the above to a proof of the theorem, we need to re-interpret this in terms of polynomials. We apply Khintchine's transference principle ([5, Theorem V.IV]) to see that (7) is satisfied infinitely often if and only if

$$|\mathbf{q} \cdot \mathbf{y} - p| < H(\mathbf{q})^{-(n+\delta'')}, \tag{8}$$

for infinitely many $\mathbf{q} \in \mathbb{Z}^n$, $p \in \mathbb{Z}$, where $\delta'' > 0$ can be explicitly bounded in terms of n and δ'. Now, $\mathbf{y}$ lies in the image of f, so that the coordinates of y consist of all monomials in the variables $(x_1, \dots x_d)$, whence any polynomial in these d variables may be expressed on the form $P(\mathbf{x}) = \mathbf{q} \cdot \mathbf{y} - p$. The coefficients of P include all the coordinates of $\mathbf{q}$ and hence $H(P) \ge H(\mathbf{q})$, so that if (3) holds for infinitely many P with $\varepsilon = \delta''$, then (8) holds for infinitely many $\mathbf{q}, p$. Since the latter condition is satisfied on a set of μ-measure zero, it follows that μ-almost all points in $\mathbb{R}^d$ are not are not k-very well approximable.

The final statement of the theorem follows immediately, at the Lebesgue measure clearly is Federer and absolutely decaying. □

Some interesting open questions present themselves at this stage. One can ask whether a vector exists which is k-very well approximable for all k. We will call such vectors very very well approximable. It is not difficult to prove that the set of k-very well approximable vectors is a dense G_δ-set, so the question of existence can be easily answered in the affirmative. However, determining the Hausdorff dimension of the set of very very well approximable vectors is an open question. When $d = 1$, it is known that the Hausdorff dimension is equal to 1 due to work of Durand [6], but the methods of that paper do not easily extend to larger values of d.

Taking the notion one step further, one can ask whether vectors $\mathbf{x} \in \mathbb{R}^d$ exist such that for some fixed $\varepsilon > 0$, for any $k \in \mathbb{N}$, there are infinitely many integer polynomials P in d variables of total degree at most k, such that

$$|P(\mathbf{x})| \le H(P)^{-(n+\varepsilon)},$$

where as usual $n = \binom{n+d}{d} - 1$, i.e. in addition to $\mathbf{x}$ being very very well approximable, we require the very very very significant improvement in the rate of approximation to be uniform in k. We will call such vectors very very very well approximable. Determining the Hausdorff dimension of the set of very very very well approximable numbers is an open problem.

2.2. **k-badly approximable points.** A point $\mathbf{x} = (x_1, \dots, x_d) \in \mathbb{R}^d$ is called k-badly approximable if there exists $C = C(k, \mathbf{x})$ such that

$$|P(\mathbf{x})| \geq CH(P)^{-n},$$

for all non-zero polynomials $P \in \mathbb{Z}[X_1, \dots, X_d]$ of total degree at most k. In other words, a point $\mathbf{x} \in \mathbb{R}^d$ is k-badly approximable if the approximation rate in Corollary 2 can be improved by at most a positive constant in the denominator. Let B_k be the set of k-badly approximable points. Note that each set B_k is a null set, which is easily deduced from the work of Beresnevich, Bernik, Kleinbock and Margulis [2]. We will now show:

Theorem 4. *Let $B \subseteq \mathbb{R}^d$ be an open ball and let $M \in \mathbb{N}$. Then*

$$\dim B \cap \bigcap_{k=1}^{M} B_k = d.$$

This statement is deduced from the work of Beresnevich [1], who proved the case $d = 1$.

Proof. Let $n_k = \binom{k+d}{d} - 1$ as before, but with the dependence on k made explicit in notation. Let $f : \mathbb{R}^d \to \mathbb{R}^{n_M}$ be given by $f(x_1, \dots, x_d) = (x_1, x_2, \dots, x_{d-1} x_d^{M-1}, x_d^M)$, with the monomials ordered in blocks of increasing total degree. Let $\mathbf{r}_k = (\frac{1}{n_k}, \dots, \frac{1}{n_k}, 0, \dots, 0) \in \mathbb{R}^{n_M}$, where the non-zero coordinates are the first n_k coordinates, so that $\mathbf{r}_k$ is a probability vector.

We define as in [1] the set of $\mathbf{r}$-approximable points for a probability vector $\mathbf{r}$ to be the set

$$\operatorname{Bad}(\mathbf{r}) = \Big\{ \mathbf{y} = (y_1, \dots, y_{n_M}) : \text{ for some } C(\mathbf{y}) > 0, \\ \max_{1 \leq i \leq n_M} \|q y_i\|^{1/r_i} \geq C(\mathbf{y}) q, \text{ for any } q \in \mathbb{N} \Big\}.$$

Here, $\|z\|$ denotes the distance to the nearest integer, and we use the convention that $z^{1/0} = 0$.

Let $1 \leq k \leq M$ be fixed and let $\mathbf{x} \in \mathbb{R}^d$ satisfy that $f(\mathbf{x}) \in \operatorname{Bad}(\mathbf{r}_k)$. From [1, Lemma 1]) it follows, that there exists a constant $C = C(k, \mathbf{x})$, such that the only integer solution $(a_0, a_1, \dots, a_{n_k})$ to the system

$$\left| a_0 + a_1 x_1 + a_2 x_2 + \dots + a_{n_k - 1} x_{d-1} x_d^{k-1} + a_{n_k} x_d^k \right| < CH^{-1}, \max_i |a_i| < H^{1/n_k}$$

is zero. Here, the choice of $\mathbf{r}_k$ and the ordering of the monomials in the function f ensure that the effect of belonging to $\operatorname{Bad}(\mathbf{r}_k)$ will only give a polynomial expression of total degree at most k. Indeed, writing out the full equivalence, we would have the first inequality unchanged, with the second being $\max_i |a_i| < H^{r_{k,i}}$, where the exponent is the i'th coordinate of $\mathbf{r}_k$. If this coordinate is 0, we are only considering polynomials where the corresponding a_i is equal to zero.

Rewriting this in terms of polynomials, for any non-zero $P \in \mathbb{Z}[X_1, \dots, X_d]$ with $H(P) < H^{1/n_k}$ and total degree at most k, we must have

$$|P(\mathbf{x})| \geq CH^{-1} > CH(P)^{-n_k}.$$

It follows that $\mathbf{x} \in B_k$, and hence $f^{-1}(\mathrm{Bad}(\mathbf{r}_k)) \subseteq B_k$. The result now follows by applying [1, Theorem 1], which implies that the Hausdorff dimension of the intersection of the sets $f^{-1}(\mathrm{Bad}(\mathbf{r}_k))$ is maximal. □

Again, an interesting open problem presents itself, namely the question of uniformity of the constant $C(k, \mathbf{x})$ in k. Is it possible to construct a vector in B_k for all k with the constant being the same for all k? And in the affirmative case, what is the Hausdorff dimension of this set? A weaker version of this question would be to ask whether there is some natural dependence of $C(k, \mathbf{x})$ on k, i.e. whether one can choose $C(k, \mathbf{x}) = C(\mathbf{x})^k$ or a similar dependence. We do not at present know the answer to these questions.

2.3. **(k, ε)-Dirichlet improvable vectors and k-singular vectors.** Let $\varepsilon > 0$. A point $\mathbf{x}$ is called (k, ε)-Dirichlet improvable if for any ε there exists a $Q_0 \in \mathbb{N}$, such that for any $Q \geq Q_0$ there exists a polynomial $P \in \mathbb{Z}[X_1, \dots, X_d]$ with total degree at most k,

$$\tilde{H}(P) \leq \varepsilon Q \quad \text{and} \quad |P(\mathbf{x})| \leq \varepsilon Q^{-n}.$$

Note that we are now using $\tilde{H}$ as a measure of the complexity of our polynomials.

In view of Theorem 1, if $\varepsilon \geq 1$, all points clearly have this property, and so the property is only of interest when $\varepsilon < 1$. A vector is called k-singular if it is (k, ε)-Dirichlet improvable for every $\varepsilon > 0$.

We will need a few additional definitions before proceeding. For a function $f : \mathbb{R}^d \to \mathbb{R}^n$, a measure μ on $\mathbb{R}^d$ and a subset $B \in \mathbb{R}^d$ with $\mu(B) > 0$, we define

$$\|f\|_{\mu,B} = \sup_{\mathbf{x} \in B \cap \operatorname{supp} \mu} |f(\mathbf{x})|.$$

Let $C, \alpha > 0$ and let $U \subseteq \mathbb{R}^d$ be open. We will say that the function f is (C, α)-good with respect to μ on U if for any ball $B \subseteq U$ with centre in $\operatorname{supp} \mu$ and any $\varepsilon > 0$,

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

We will say that a measure μ on $\mathbb{R}^d$ is k-friendly if it is Federer, non-planar and the function $f : \mathbb{R}^d \to \mathbb{R}^n$ given by $f(x_1, \dots, x_d) = (x_1, x_2, \dots, x_{d-1} x_d^{k-1}, x_d^k)$ is (C, α)-good with respect to μ on $\mathbb{R}^d$ for some $C, \alpha > 0$.

We have

Theorem 5. *Let μ be a k-friendly measure on $\mathbb{R}^d$. Then there is an $\varepsilon_0 = \varepsilon_0(d, \mu)$ such that the set of (k, ε)-Dirichlet improvable points has measure zero for any $\varepsilon < \varepsilon_0$. In particular, the set of k-singular vector has measure zero.*

In the case when $d = 1$, $k \geq 2$ and μ being the Lebesgue measure on $\mathbb{R}$, the result is immediate from work of Bugeaud [3, Theorem 7], in which an explicit value of ε is given, namely $\varepsilon = 2^{-3k-3}$. Our proof is non-effective and relies on [9, Theorem 1.5].

Proof. Under the assumption on the measure μ, [9, Theorem 1.5] implies the existence of an $\varepsilon_0 > 0$ such that for all $\tilde{\varepsilon} < \varepsilon_0$

$$f_* \mu(\mathrm{DI}_{\tilde{\varepsilon}}(\mathcal{T})) = 0 \text{ for any unbounded } \mathcal{T} \subseteq \mathfrak{a}_+.$$

Here, f is the usual function $f(x_1,\dots,x_d) = (x_1, x_2, \dots, x_{d-1}x_d^{k-1}, x_d^k)$, $\mathfrak{a}^+$ denotes the set of $(n+1)$-tuples of $(t_0, t_1, \dots, t_n)$ such that $t_0 = \sum_{i=1}^n t_i$, $t_i > 0$ for each i, and $\mathrm{DI}_{\tilde{\varepsilon}}(\mathcal{T})$ denotes the set of vectors $\mathbf{y} = (y_1,\dots,y_n) \in \mathbb{R}^n$ for which there is a T_0 such that for any $t \in \mathcal{T}$ with $\|t\| \geq T_0$, the system of inequalities

$$\begin{cases} |\mathbf{q}\cdot\mathbf{y} - p| < \tilde{\varepsilon}e^{-t_0} \\ |q_i| < \tilde{\varepsilon}e^{t_i} \qquad\qquad i = 1,\dots,n, \end{cases}$$

has infinitely many non-trivial integer solutions $(\mathbf{q}, p) = (q_1,\dots,q_n,p) \in \mathbb{Z}^{n+1} \setminus \{0\}$.

Our result follows by specialising the above property. Indeed, we apply this to $\varepsilon = \tilde{\varepsilon}^{n+1} < \varepsilon_0^{n+1}$ and the central ray in $\mathfrak{a}^+$,

$$\mathcal{T} = \left\{ \left(t, \frac{t}{n}, \dots, \frac{t}{n}\right) : t = \log\left(\frac{Q}{\tilde{\varepsilon}}\right) n, Q \geq [\varepsilon_0] + 1, Q \in \mathbb{N} \right\}.$$

The measure $f_*\mu$ is the pushforward under f of the k-friendly measure μ. It follows that the set of $\mathbf{x} \in \mathbb{R}^d$ for which their image under f is in $\mathrm{DI}_{\tilde{\varepsilon}}(\mathcal{T})$ is of measure zero for all $\tilde{\varepsilon} < \varepsilon_0^{n+1}$. From the definition of $\mathrm{DI}_{\tilde{\varepsilon}}$ and the choice of $\mathfrak{a}^+$ and $\mathcal{T}$, $f(\mathbf{x}) \in \mathrm{DI}_{\tilde{\varepsilon}}$ if and only if there is a $Q_0 \geq \max\{[\varepsilon_0] + 1, \tilde{\varepsilon}e^{T_0/n}\}$ such that for $Q > Q_0$ there exists $q_0, q_1, \dots, q_n \in \mathbb{Z}$ with $\max_{1\leq i\leq n} |q_i| < \tilde{\varepsilon}e^{t/n} = Q$ such that

$$|(q_1, \dots q_n) \cdot f(\mathbf{x}) + q_0| < \tilde{\varepsilon}e^{-t} = \varepsilon Q^{-n}.$$

Reinterpreting the right hand side of the above as a polynomial expression in $\mathbf{x}$, this recovers the exact definition of $\mathbf{x}$ being $(k, \varepsilon^{1/(n+1)})$-Dirichlet improvable. □

Note that the proof in fact yields a stronger statement. Namely, by adjusting the choice of $\mathfrak{a}^+$, we could have put different weights on the coefficients of the approximating polynomials, thus obtaining the same result, but with a non-standard (weighted) height of the polynomial.

As with the preceding results, some open problems occur. We do not at present know if there exist a vector $\mathbf{x}$, for which there are positive numbers $\varepsilon_k > 0$, such that $\mathbf{x} \in \mathrm{DI}(k, \varepsilon_k)$. If this is the case, determining the Hausdorff dimension of the set of such vectors is another open problen. Additionally, the same questions can be asked if we require ε to be independent of k, i.e. if we ask for the existence of a vector $\mathbf{x} \in \mathrm{DI}(k, \varepsilon)$ for all k.

Let us now say that $\mathbf{x} \in \mathbb{R}^d$ is k-algebraic if there exists a nontrivial polynomial $P \in \mathbb{Z}[X_1, \dots, X_d]$ of degree at most k, such that $P(\mathbf{x}) = 0$. It is clear that if $\mathbf{x}$ is k-algebraic, then it is k-singular. In light of Theorem 5, it is natural to inquire whether all k-singular points are k-algebraic. In this direction we have:

Theorem 6. *For $d \geq 2$, for any $k \geq 1$, there exists a k-singular point in $\mathbb{R}^d$ which is not k-algebraic.*

The proof relies on results of [8]. We do not know whether the conclusion of Theorem 6 is valid for $d = 1$.

Proof. Once more, for a fixed k, we take f as in the proof of Theorem 5. In the notation of [8], it is clear that $\mathbf{x} \in \mathbb{R}^d$ is k-singular if $f(\mathbf{x}) \in \mathrm{Sing}(\mathbf{n})$. Also $f(\mathbf{x})$ is totally irrational in the notation of [8] if and only if $\mathbf{x}$ is k-algebraic.

Since the image of f is a d-dimensional nondegenerate analytic submanifold of $\mathbb{R}^n$, for $d \geq 2$ we can apply [8, Theorem 1.2] to conclude that the intersection of $f(\mathbb{R}^d)$ with $\mathrm{Sing}(\mathbf{n})$ contains a totally irrational point. □

Theorem 5 does not give an explicit value of ε_0, and indeed the value depends on the measure μ. However we can at least push ε_0 to the limit $\varepsilon_0 \nearrow 1$ in the case when μ is the Lebesgue measure on $\mathbb{R}^d$ to obtain a result on the k-singular vectors.

Theorem 7. *For any d, the set of $\mathbf{x}$ which are (k, ε)-Dirichlet improvable for some $\varepsilon < 1$ and some k, has Lebesgue measure zero.*

The proof relies on the work of Shah [15].

Proof. This is a direct consequence of [15, Corollary 1.4], where the set $\mathcal{N}$ is chosen to be the diagonal $\mathcal{N} = \{(N, \dots, N) : N \in \mathbb{N}\}$. □

Note that once again, the result of Shah gives a stronger result in the sense that we may take a non-standard height as in the preceding case and retain the conclusion.

2.4. **Algebraic vectors.** Our final result, which is again a corollary of known results, is an analogue of Roth's Theorem [12], which states that algebraic numbers are not very well approximable. Schmidt's Subspace Theorem, see e.g. [13], provides a higher dimensional analogue of this result, and it is this theorem we will apply. We will say that a vector $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_d) \in \mathbb{R}^d$ is algebraic of total degree k if there is a polynomial $P_{\boldsymbol{\alpha}} \in \mathbb{Z}[X_1, \dots, X_d]$ of total degree k with $P_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}) = 0$ and if no polynomial of lower total degree vanishes at $\boldsymbol{\alpha}$.

Theorem 8. *Let $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_d) \in \mathbb{R}^d$ be an algebraic d-vector of total degree more than k. Then for any $\varepsilon > 0$ there are only finitely many non-zero polynomials $P \in \mathbb{Z}[X_1, \dots, X_d]$ of total degree at most k with*

$$|P(\boldsymbol{\alpha})| < H(P)^{-(n+\varepsilon)}.$$

Proof. Since $\boldsymbol{\alpha}$ in not algebraic of total degree at most k, by definition it follows that the numbers $1, \alpha_1, \alpha_2, \dots, \alpha_{d-1}\alpha_d^{k-1}, \alpha_d^k$ are algebraically independent over $\mathbb{Q}$. From a corollary to Schmidt's Subspace Theorem, [13, Chapter VI Corollary 1E], it follows that there are only finitely many non-zero integer solutions $(q_0, \dots, q_n)$ to

$$\left|q_0 + q_1\alpha_1 + q_2\alpha_2 + \dots + q_{n-1}\alpha_{d-1}\alpha_d^{k-1} + q_n\alpha_d^k\right| < \left(\max_{1\le i\le n} |q_i|\right)^{-(n+\varepsilon)}.$$

This immediately implies the result.

□

References

1. V. Beresnevich, *Badly approximable points on manifolds*, Invent. Math. **202** (2015), no. 3, 1199–1240. MR 3425389
2. V. V. Beresnevich, V. I. Bernik, D. Y. Kleinbock, and G. A. Margulis, *Metric Diophantine approximation: the Khintchine-Groshev theorem for nondegenerate manifolds*, Mosc. Math. J. **2** (2002), no. 2, 203–225, Dedicated to Yuri I. Manin on the occasion of his 65th birthday. MR 1944505 (2004b:11107)
3. Y. Bugeaud, *Approximation by algebraic integers and Hausdorff dimension*, J. London Math. Soc. (2) **65** (2002), no. 3, 547–559. MR 1895732 (2003d:11110)
4. Y. Bugeaud, *Approximation by algebraic numbers*, Cambridge University Press, 2004.
5. J. W. S. Cassels, *An introduction to diophantine approximation*, Cambridge University Press, 1957.
6. A. Durand, *Large intersection properties in Diophantine approximation and dynamical systems*, J. Lond. Math. Soc. (2) **79** (2009), no. 2, 377–398. MR 2496520 (2010b:11084)
7. D. Y. Kleinbock, E. Lindenstrauss, and B. Weiss, *On fractal measures and Diophantine approximation*, Selecta Math. (N.S.) **10** (2004), no. 4, 479–523. MR 2134453 (2006g:11151)

8. D. Y. Kleinbock and B. Weiss, *Friendly measures, homogeneous flows and singular vectors*, **385** (2005), 281–292. MR 2180240 (2006f:11084)
9. ———, *Dirichlet's theorem on Diophantine approximation and homogeneous flows*, J. Mod. Dyn. **2** (2008), no. 1, 43–62. MR 2366229 (2008k:11078)
10. J. F. Koksma, *Über die mahlersche klasseneinteilung der transzendenten zahlen und die approximation komplexer zahlen durch algebraische zahlen*, Monatsh. Math. Phys. **48** (1939), 176–189.
11. K. Mahler, *Zur approximation der exponentialfunktionen und des logarithmus, i, ii*, J. Reine Angew. Math. **166** (1932), 118–150.
12. K. F. Roth, *Rational approximations to algebraic numbers*, Mathematika **2** (1955), 1–20; corrigendum, 168. MR 0072182 (17,242d)
13. W. M. Schmidt, *Diophantine approximation*, Lecture Notes in Mathematics, vol. 785, Springer, Berlin, 1980. MR 568710 (81j:10038)
14. ———, *Mahler and Koksma classification of points in* $\mathbb{R}^n$ *and* $\mathbb{C}^n$, Funct. Approx. Comment. Math. **35** (2006), 307–319. MR 2271620 (2008c:11105)
15. N. A. Shah, *Expanding translates of curves and Dirichlet-Minkowski theorem on linear forms*, J. Amer. Math. Soc. **23** (2010), no. 2, 563–589. MR 2601043 (2011g:11137)
16. K. R. Yu, *A generalization of Mahler's classification to several variables*, J. Reine Angew. Math. **377** (1987), 113–126. MR 887404 (88h:11049)

S. KRISTENSEN, Department of Mathematics, Aarhus University, Ny Munkegade 118, DK-8000 Aarhus C, Denmark
E-mail address: `sik@math.au.dk`

S. H. PEDERSEN, Department of Mathematics, Aarhus University, Ny Munkegade 118, DK-8000 Aarhus C, Denmark
E-mail address: `steffenh@math.au.dk`

B. WEISS, Department of Mathematics, Tel Aviv University, Tel-Aviv, 69978 Israel
E-mail address: `barakw@post.tau.ac.il`