JOURNAL OF APPROXIMATION THEORY **69**, 231–249 (1992)

Dimension and Geometry of Sets Defined by Polynomial Inequalities

HANS WALLIN AND PETER WINGREN

Department of Mathematics, University of Umeå, S-901 87 Umeå, Sweden

Communicated by Paul Nevai

Received January 28, 1991; revised May 15, 1991

Markov's inequality for the maxima of the derivatives of polynomials over cubes is replaced by an inequality where the cubes are changed to certain cubes intersected by a given subset F of $\mathbb{R}^n$. This new inequality is true for certain sets F and false for others. We are interested in the sets F for which this inequality is true and we prove that these sets must have positive Hausdorff dimension. Our inequality is not true if F is the closure of a domain with an outgoing cusp. We introduce a generalized inequality which holds for these sets and prove that this new inequality allows sets F with Hausdorff dimension zero.

INTRODUCTION

The inequalities of Bernstein and Markov for the derivative of a polynomial are of importance in the proof of inverse theorems in approximation theory. The inequality of A. A. Markov states that if P is an algebraic polynomial in one variable of degree k then

$$\max_{-1 \leqslant x \leqslant 1} |P'(x)| \leqslant k^2 \max_{-1 \leqslant x \leqslant 1} |P(x)|.$$

For algebraic polynomials P in n real variables of total degree k and with the interval $[-1, 1]$ replaced by any n-dimensional cube Q with side of length 2δ, $0 < \delta$, Markov's inequality becomes

$$\max_{Q} |\nabla P| \leqslant \frac{c}{\delta} \max_{Q} |P|, \tag{1}$$

with $c = k^2 \sqrt{n}$.

If F is a given closed subset of $\mathbb{R}^n$ there are different possibilities of replacing (1) by an inequality stating that Markov's inequality is, in some sense, valid on F. This new inequality may be true for some sets F and false for others, and, consequently, it gives a condition on the set F. Properly

0021-9045/92 $5.00

chosen this inequality serves a purpose for approximation of functions defined on F similar to the original Bernstein and Markov inequalities on an interval.

One such possibility is to replace Q by $Q \cap F$ in (1) and to assume that this new inequality is true for all cubes Q with side $2\delta \leqslant 1$ and center in F, and for all polynomials P, with a constant $c = c(F, n, k)$ depending only on F, n, and the degree k of P. This gives a condition on F which we express by saying that F preserves Markov's inequality. This condition on F is important in the study of polynomial interpolation and function spaces on F, in particular smoothness of functions defined on F by means of local polynomial approximation (see [3, 8–10] and Section 1.1). The condition is studied in Section 1; see Section 1.1 for the precise definition and for a geometrical characterization of these sets. It turns out that the class of sets preserving Markov's inequality contains many fractal sets including generalized Cantor sets of arbitrarily small, but positive, Hausdorff dimension. Our main result is that, on the other hand, these sets cannot be too small; in fact they must have positive Hausdorff dimension (Corollary 1 in Section 1.4). As a preliminary, in Section 1.1 we have collected some known results about sets preserving Markov's inequality, and some motivating background material; in Section 1.2 we give a result on the Hausdorff dimension of certain sets of a generalized Cantor type; and in Section 1.3 we give a suitable geometrical characterization of sets preserving Markov's inequality.

A set which is the closure of a domain with an outgoing cusp does not preserve Markov's inequality in the sense of Section 1. This is part of the motivation for Section 2, where we study a generalized version of Markov's inequality on F, a version which allows cusp domains. The cusp domain is studied in Section 2.1 where we also give some further motivation. The generalized Markov inequality is treated in Section 2.2. It is proved in Section 2.3 that sets preserving this generalized Markov inequality may have Hausdorff dimension zero (Proposition 11); this should be compared to our main result, Corollary 1.

Finally we want to mention that another often studied version of (1) for compact sets F is obtained by replacing Q by F in (1) and c/δ by a constant which is allowed to grow polynomially in the degree of P. This version of Markov's inequality is important when studying smoothness of functions on F by means of polynomial approximation over the whole of F [12, Theorem 3.3 and 4.2], as well as for a number of other problems [6].

Notation. $\mathbb{R}^n$ is the n-dimensional Euclidean space with points $x = (x_1, ..., x_n)$ and the usual norm $|x|$. F is a closed non-empty subset of $\mathbb{R}^n$. $B(x, r)$ is the closed n-dimensional ball with center x and radius r. Q

is an n-dimensional cube with sides parallel to the coordinate axes. The diameter of $A \subset \mathbb{R}^n$ is denoted diam A. The gradient is denoted by ∇ and the maximum norm of f over A by $\|f\|_A$. $\mathscr{P}_k$ denotes the set of algebraic polynomials in n real variables of total degree at most k.

1. Sets Preserving Markov's Inequality

1.1. We start with the following formal definition where we use balls B instead of the cubes Q used in the introduction. However, this gives an equivalent condition (see Remark 1 below).

Definition 1. A closed non-empty subset F of $\mathbb{R}^n$ *preserves Markov's inequality* if for every positive integer k there exists a constant $c = c(F, n, k)$, depending only on F, n, and k, such that

$$\|\nabla P\|_{B \cap F} \leqslant \frac{c}{r} \|P\|_{B \cap F}, \tag{2}$$

for all $P \in \mathscr{P}_k$ and all balls $B = B(x_0, r)$ with $x_0 \in F$ and $0 < r \leqslant 1$. We call (2) *Markov's inequality on F*.

As mentioned in the introduction (2) is important in the study of smoothness properties of functions defined on F. We refer to [3] for a detailed study and here we mention only the following result: Let F be a set preserving Markov's inequality, let α be a positive number (the smoothness index), and let f be a bounded function on F. Then (see [3, p. 72]) f is the restriction to F of a function in the Lipschitz space $\Lambda_\alpha(\mathbb{R}^n)$ (see [3, p. 2] for the definition of $\Lambda_\alpha(\mathbb{R}^n)$) if and only if there exists a constant c_1 so that, for every cube Q with center in F and side of length $\delta \leqslant 1$, there exists a polynomial P having degree at most equal to the integer part of α such that

$$\|f - P\|_{Q \cap F} \leqslant c_1 \delta^\alpha.$$

The connection between Markov's inequality and polynomial interpolation on F is explained in [8–10].

We refer to [3, pp. 34–40] and to [8, Sect. 1], for the proof of Propositions 1–4 below.

Proposition 1. *F preserves Markov's inequality if and only if for every positive integer k there exists a constant $c_1 = c_1(F, n, k)$ such that for all balls*

$B=B(x_0, r)$, $x_0 \in F$, $0<r\leqslant 1$, and all polynomials $P(x)=\sum_{|j|\leqslant k} a_j(x-x_0)^j$ of degree at most k, we have

$$\sum_{|j|\leqslant k} |a_j|\, r^{|j|} \leqslant c_1 \|P\|_{B\cap F}.$$

Remark 1. By using this proposition—or, alternatively, to argue as in Section 2.2 below—it is possible to realize that in Definition 1 we may change the condition $0<r\leqslant 1$ to $0<r\leqslant r_0$ for any constant $r_0>0$, and the balls B to cubes Q with center in F and side at most δ_0 where $\delta_0>0$ is any constant, without changing the class of sets F preserving Markov's inequality.

By using Proposition 1 or Markov's inequality in $\mathbb{R}^n$ one can also prove the following proposition, which is useful for instance when studying polynomial interpolation (see [10]).

PROPOSITION 2. *F preserves Markov's inequality if and only if for every positive integer k there exists a constant $c_1=c_1(n, F, k)$ so that*

$$\|P\|_B \leqslant c_1 \|P\|_{B\cap F} \tag{3}$$

for all $P\in\mathscr{P}_k$ and all $B=B(x_0, r)$, $x_0\in F$, $0<r\leqslant 1$.

In the right-hand member of (3) it is, for each k, possible to replace the maximum of $|P|$ over $B\cap F$ by the maximum of $|P|$ over a finite k-unisolvent subset of $B\cap F$, where the subset is independent of P (see [10, Introduction]). This gives a link between (2) and polynomial interpolation. Proposition 1 plays a crucial role in proving the somewhat surprising fact that it is enough to assume that (1) holds for $k=1$:

PROPOSITION 3. *If (2) or (3) holds for all polynomials P of degree 1 with a constant $c=c(F, n)$, for all $B=B(x_0, r)$, $x_0\in F$, $0<r\leqslant 1$, then F preserves Markov's inequality.*

Because of Proposition 3 it is possible to give the following geometric characterization of sets preserving Markov's inequality; a related geometric characterization which may also be used to prove Proposition 4 will be given in Proposition 7 in Section 1.3.

PROPOSITION 4. *F preserves Markov's inequality if and only if there exists an $\varepsilon>0$, so that for every ball $B=B(x_0, r)$, $x_0\in F$, $0<r\leqslant 1$, and every band S of type $S:=\{x\in\mathbb{R}^n\colon |b\cdot(x-x_0)|<\varepsilon r\}$, $b\in\mathbb{R}^n$, $|b|=1$ (see Fig. 1),*

$$F\cap(B\setminus S)\neq\varnothing.$$

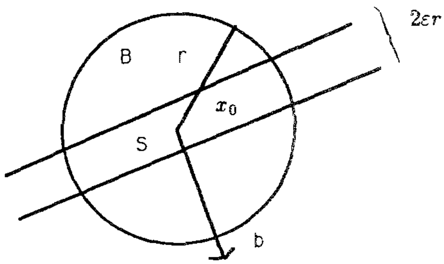

FIGURE 1

To sum up, we have in principle three methods to characterize sets preserving Markov's inequality: an algebraic given by Definition 1, a geometric given by Proposition 4, and the one in terms of polynomial interpolation mentioned after Proposition 2.

The geometric characterization in Proposition 4 means that a set F preserving Markov's inequality may not be too flat anywhere. For instance, a subset of an $(n-1)$-dimensional affine subspace of $\mathbb{R}^n$ or of the boundary of an n-dimensional ball in $\mathbb{R}^n$ does not preserve Markov's inequality. Proposition 4 may also be used to give several examples of sets preserving Markov's inequality. Such examples are the closure of an open set in $\mathbb{R}^n$ with Lipschitz boundary or of an (ε, δ)-domain. Further examples are the ordinary Cantor set, von Koch's curve and, in fact, a lot of other fractals:

PROPOSITION 5 [7, Theorem 1]. *A geometrically self-similar set* [7, Sect. 2], *which is not a subset of an* $(n-1)$*-dimensional affine subspace of* $\mathbb{R}^n$, *preserves Markov's inequality.*

We see that there are a lot of sets preserving Markov's inequality, including "small" sets like the ordinary Cantor set. The question comes up: How small can a closed, non-empty set F preserving Markov's inequality be? We see, for instance from Proposition 4, that F cannot have isolated points, i.e., F must be a perfect set and, consequently, it must be non-denumerable. On the other hand it is straightforward to construct sets F preserving Markov's inequality having Hausdorff dimension less than any prescribed positive number. In fact, for $n=1$ it is enough to take F as a generalized Cantor set where we do as in the usual Cantor construction, dividing each interval into three parts, except that we always let the length of the interval in the middle consist of a fixed, prescribed proportion of the length of the interval which we divide. By making this proportion close to one we get a generalized Cantor set F with Hausdorff dimension as close to zero as we wish (see Section 1.2), and, by Proposition 4, we may check that F preserves Markov's inequality. By working with Cartesian products of generalized Cantor sets we may extend this example to higher dimensions. We summarize this discussion (see also Remark 3 in Section 1.2):

PROPOSITION 6. *There exists a generalized Cantor set F in $\mathbb{R}^n$ preserving Markov's inequality and having Hausdorff dimension less than any prescribed positive number.*

There is a close connection between nowhere dense sets preserving Markov's inequality and sets of generalized Cantor type. We illustrate that for $n=1$ and assume that F is a closed, nowhere dense subset of $[0, 1]$ preserving Markov's inequality and containing 0 and 1. We shall see that F is a generalized Cantor set in some sense. We start from $[0, 1]$. The set $[0, 1]\setminus F$ is a union of open intervals. Let I be one of the largest of these. The endpoints of I are different from 0 and 1 since these points belong to F and F preserves Markov's inequality. We remove I from $[0, 1]$ and repeat the process on each of the two remaining intervals and remove two new intervals, and so on. After infinitely many steps we get F written as a generalized Cantor set.

By a variation of this construction we shall in Section 1.3 answer one question about the size of sets preserving Markov's inequality by showing that such a set always must contain a set of generalized Cantor type having positive Hausdorff dimension. As a preparation we estimate the Hausdorff dimension of sets of Cantor type in Section 1.2.

1.2. The Hausdorff dimension of generalized Cantor sets is determined for instance in [4, 5]. The theorems in those papers, however, deal with more general situations than the one we are interested in below and their proofs are relatively comprehensive. We prefer, because of that, to give a short proof which is easy to follow and adapted to our situation as described in Theorem 1.

First we introduce some notation. If E is a subset of $\mathbb{R}^n$, $s \geqslant 0$, and $0 < \delta \leqslant \infty$, we define

$$H^s_\delta(E) := \inf \sum_i (\operatorname{diam} U_i)^s,$$

where the infimum is taken over all countable coverings of E by sets U_i with $\operatorname{diam} U_i \leqslant \delta$. When δ decreases to zero $H^s_\delta(E)$ increases to a limit, $H^s(E)$, finite or infinite, the *s-dimensional Hausdorff measure of E*. The *Hausdorff dimension of E*, $\dim_H(E)$, is either 0 or the unique positive number s_0 such that $H^s(E)$ is infinite if $s < s_0$ and zero if $s > s_0$.

Next we turn to the construction of our generalized Cantor set E in $\mathbb{R}^n$.

We define E by means of families $\mathscr{B}_p$, $p=0, 1, 2, \ldots$, of closed n-dimensional balls by putting

$$E := \bigcap_{p=0}^{\infty} \bigcup_{B \in \mathscr{B}_p} B. \tag{4}$$

The families $\mathscr{B}_p$ are chosen inductively as follows starting from given numbers r and ρ, $r>0$, $0<\rho<1$, where ρ is assumed to be small enough:

(i) Every $B \in \mathscr{B}_p$ has radius $\rho_p := r \cdot \rho^p$.

(ii) $\mathscr{B}_0$ consists of exactly one ball.

(iii) Suppose $\mathscr{B}_p$ has been chosen. Chose $n+1$ balls with pairwise disjoint interior in every $B \in \mathscr{B}_p$. We define $\mathscr{B}_{p+1}$ as the family of all balls obtained in this way from balls in $\mathscr{B}_p$.

It follows that $\mathscr{B}_p$ consists of $(n+1)^p$ balls. We observe that the construction is possible if ρ is small enough and that E, defined by (4), has points in the interior of B for any B in any $\mathscr{B}_p$.

THEOREM 1. *Let E be the generalized Cantor set constructed above and put*

$$s := \log(n+1)/\log \frac{1}{\rho}.$$

Then there exists a positive constant M depending only on n such that

$$\frac{(2r)^s}{M} \leqslant H^s(E) \leqslant (2r)^s.$$

In particular

$$\dim_H(E) = \log(n+1) \Big/ \log \frac{1}{\rho}.$$

Proof. Take $\delta>0$. As E is covered by the balls in $\mathscr{B}_p$, for every p, we get an estimate from above,

$$H^s_\delta(E) \leqslant (n+1)^p \, (2r\rho^p)^s = (2r)^s. \tag{5}$$

To get an estimate from below we proceed in three steps.

(I) We start with an open covering of E with sets U_i, $i=1, 2, \ldots$, of diameter at most δ. Since E is compact there exists a finite subcover U_i', $i=1, 2, \ldots, N$, such that every U_i' intersects E.

(II) We observe that

$$U_i' \cap E \subset \bigcup_{B \in \mathscr{B}_p} B, \qquad \text{for every } U_i' \text{ and } p.$$

Pick, for any chosen U_i', the unique family $\mathscr{B}_p$ such that $2\rho_p < \operatorname{diam} U_i' \leqslant 2\rho_{p-1}$, where ρ_p, $p \geqslant 0$, were defined in the constructon of $\mathscr{B}_p$; we put $\rho_{-1} := \infty$. Since all the balls in $\mathscr{B}_p$ have the same radius there exists a number $M = M(n)$ which is *independent of* i *and* δ and gives an upper bound of the number of balls in $\mathscr{B}_p$ intersecting U_i'. Denote the intersecting balls by B_{ij}, $1 \leqslant j \leqslant j(i) \leqslant M$. Hence, $\beta := \{B_{ij}\}$, $1 \leqslant j \leqslant j(i)$, $1 \leqslant i \leqslant N$, is a finite covering of E and

$$(\operatorname{diam} U_i')^s > (2\rho_p)^s = \frac{1}{j(i)} \sum_{j=1}^{j(i)} (\operatorname{diam} B_{ij})^s \geqslant \frac{1}{M} \sum_{j=1}^{j(i)} (\operatorname{diam} B_{ij})^s.$$

(III) Remove every ball in β which is a subset of some other ball in β and assume that the smallest of the remaining balls in β has radius ρ_m. Then we successively replace every remaining ball in β having radius larger than ρ_m by balls in $\mathscr{B}_m$ so that we obtain a covering of E by balls in $\mathscr{B}_m$ only. In this process we observe that if $B \in \mathscr{B}_\nu$ and $B_1, B_2, ..., B_{n+1} \in \mathscr{B}_{\nu+1}$, then

$$(\operatorname{diam} B)^s = \sum_{i=1}^{n+1} (\operatorname{diam} B_i)^s.$$

This gives a covering of E consisting of *all* the balls in $\mathscr{B}_m$, because, by the construction of E, the interior of every $B \in \mathscr{B}_m$ contains points of E. We now get the estimate

$$\begin{aligned}\sum_i (\operatorname{diam} U_i)^s &\underset{(\mathrm{I})}{\geqslant} \sum_1^N (\operatorname{diam} U_i')^s \underset{(\mathrm{II})}{\geqslant} \sum_{i=1}^N \frac{1}{M} \sum_{j=1}^{j(i)} (\operatorname{diam} B_{ij})^s \\ &\underset{(\mathrm{III})}{\geqslant} \frac{1}{M} \sum_{B \in \mathscr{B}_m} (\operatorname{diam} B)^s = \frac{1}{M} (2r\rho^m)^s \cdot (n+1)^m = \frac{(2r)^s}{M},\end{aligned}$$

which gives

$$H_\delta^s(E) \geqslant \frac{(2r)^s}{M}.$$

If we combine this with (5) and let δ tend to zero, we get the theorem.

Remark 2. It follows from the proof that Theorem 1 is true with $H^s(E)$ replaced by $H_\infty^s(E)$.

Remark 3. By choosing ρ small we can get Proposition 6 as a corollary to Theorem 1.

1.3. In the proof of our main theorem (Theorem 2 in Section 1.4) we need a modification (Proposition 7) of the geometric characterization in Proposition 4 in Section 1.1. Similar modifications have been proved in [2, 9].

For an $F \subset \mathbb{R}^n$ preserving Markov's inequality we are interested in the best constant $c = c(F, n, k)$ in (2) for $k = 1$ and refer to it as the best constant in Markov's inequality on F for first degree polynomials. Let us call this constant $c_1 = c_1(F, n)$. We are also interested in the best constant c in (2) for first degree polynomials which are zero at x_0, i.e., such that for each $B = B(x_0, r)$, $x_0 \in F$, $0 < r \leqslant 1$, (2) holds for all $P \in \mathscr{P}_1$ with $P(x_0) = 0$. We refer to that c as *the best constant in Markov's inequality on F for first degree polynomials which vanish at the center*. Let us denote this constant by $c_0 = c_0(F, n)$. We claim that

$$c_0 \leqslant c_1 \leqslant 2c_0.$$

In fact, the left hand inequality follows from the definition of c_0 and c_1 and the right hand inequality from the following calculation. Let x_0 and r be given, $x_0 \in F$, $0 < r \leqslant 1$, and let $P(x) = \nabla P \cdot (x - x_0) + P(x_0)$ be an arbitrary first degree polynomial. Then P_1 defined by $P_1(x) = P(x) - P(x_0)$ is zero at x_0 and we get

$$|\nabla P| = |\nabla P_1| \leqslant \frac{c_0}{r} \|P_1\|_{B \cap F} \leqslant \frac{2c_0}{r} \|P\|_{B \cap F},$$

where the last inequality follows from the triangle inequality. From this chain of inequalities we see that $c_1 \leqslant 2c_0$, proving our claim.

We remark that a consequence of our discussion and Proposition 3 is that F preserves Markov's inequality if F is any closed non-empty subset of $\mathbb{R}^n$ such that, for some constant c, (2) holds for every $B = B(x_0, r)$, $x_0 \in F$, $0 < r \leqslant 1$, and every $P \in \mathscr{P}_1$ with $P(x_0) = 0$.

PROPOSITION 7. *Let F be a closed non-empty subset of $\mathbb{R}^n$ and c_0 a positive constant. The following two conditions are equivalent.*

(I) *F preserves Markov's inequality and the best constant in Markov's inequality on F for first degree polynomials which vanish at the center, is c_0.*

(II) *For every ball $B = B(x_0, r)$, $x_0 \in F$, $0 < r \leqslant 1$, and every $(n-1)$-dimensional affine subspace H of $\mathbb{R}^n$ containing x_0, there exists a point in $B \cap F$ at distance larger than or equal to r/c_0 from H.*

Proof. We first assume (I). Given B and H as in (II) let b be a unit normal to H and introduce $P(x) = b \cdot (x - x_0)$. By (I) we get

$$1 = |b| = |\nabla P| \leqslant \frac{c_0}{r} \max_{x \in B \cap F} |b \cdot (x - x_0)|.$$

Hence, $|b \cdot (x - x_0)|$, which means the distance from x to H, is larger than or equal to r/c_0 for some $x \in B \cap F$, proving (II); in particular we see that $c_0 \geqslant 1$.

To prove the other half of the theorem we just follow the discussion in the other direction. We start with $P(x) = b \cdot (x - x_0)$, which we may normalize by assuming that $|b| = 1$, then choose H with normal b, and we see that the geometric condition in (II) gives the Markov inequality required in (I).

1.4. We can now state and prove our main theorem on the Hausdorff measure of sets preserving Markov's inequality (2).

THEOREM 2. *Assume that $F \subset \mathbb{R}^n$ preserves Markov's inequality. Let c_0 be the best constant in Markov's inequality on F for first degree polynomials which vanish at the center (defined in Section 1.3), and introduce*

$$s := \frac{\log(n+1)}{\log(1 + 2c_0)}.$$

Then there exists a constant $c_2 > 0$ depending only on n such that

$$\frac{H^s(F \cap B(x_0, r))}{(2r)^s} \geqslant c_2,$$

for all $x_0 \in F$ and all $r \in (0, 1]$.

As a corollary we get our main result:

COROLLARY 1. *Every closed non-empty set preserving Markov's inequality has positive Hausdorff dimension.*

In the proof of Theorem 2 we shall use Proposition 7 to show the existence of a subset E of F in Theorem 2 of generalized Cantor type constructed by means of families $\mathscr{B}_p$ of balls as in Section 1.2. Theorem 2 will then follow from Theorem 1.

Proof of Theorem 2. (1) Take $x_0 \in F$, $r \in (0, 1]$ and $B = B(x_0, r)$. By Proposition 7 we can find $n+1$ affinely independent points $y_0, y_1, ..., y_n$ in $B \cap F$ such that

(i) $y_0 = x_0$, and

(ii) if $y_0, \ldots, y_j$ are chosen and $j < n$ we first choose an $(n-1)$-dimensional affine subspace H_j of $\mathbb{R}^n$ containing $y_0, \ldots, y_j$ and then y_{j+1} in $B \cap F$ at distance at least r/c_0 from H_j.

Our choice means that $|y_{j+1} - y_i|$ for $i \leqslant j$ is larger than or equal to the distance from y_{j+1} to H_j which is at least r/c_0. We conclude that

$$|y_j - y_i| \geqslant r/c_0 \qquad \text{if} \quad i \neq j.$$

(2) Given $B = B(x_0, r)$ take $B(x_0, \sigma)$, where $\sigma < r$ is chosen below, and choose $y_0, \ldots, y_n$ as in Step 1 but with B replaced by $B(x_0, \sigma)$, i.e., r by σ. This means that $y_j \in F \cap B(x_0, \sigma)$ and $|y_j - y_i| \geqslant \sigma/c_0$ if $i \neq j$. We now take balls $B(y_j, \rho_1)$, $j = 0, \ldots, n$, where

$$\rho_1 = \frac{\sigma}{2c_0}.$$

Consequently, these new balls have pairwise disjoint interior, and they are all subsets of $B(x_0, r)$ if

$$\rho_1 = r - \sigma.$$

We choose σ and ρ_1 so that the two last conditions hold, which gives

$$\rho_1 = \frac{r}{1 + 2c_0}.$$

(3) Now we have the machinery needed to use Theorem 1 in Section 1.2. The families $\mathscr{B}_p$, $p = 0, 1, \ldots,$ of balls with radius $\rho_p = r\rho^p$ are constructed as follows with $\rho = (1 + 2c_0)^{-1}$. The first family, $\mathscr{B}_0$, consists of the single ball $B = B(x_0, r)$, and the second family, $\mathscr{B}_1$, of the balls $B(y_j, \rho_1)$ constructed in Step 2. By repeating the construction in Step 2 on each of the balls in $\mathscr{B}_1$ we get $\mathscr{B}_2$, and so on. From the families $\mathscr{B}_p$ we get a generalized Cantor set E as in Section 1.2 and it follows that $E \subset F \cap B$ from the construction and the fact that F is closed. Finally, Theorem 2 follows from Theorem 1 with $c_2 = 1/M$.

Remark 4. It follows from Remark 2 in Section 1.2 that in Theorem 2 we may replace $H^s(F \cap B(x_0, r))$ by $H^s_\infty(F \cap B(x_0, r))$. This should be compared to Theorem 3 in Section 1.5.

Remark 5. It appears from the proof of Theorem 2 that F locally always has a subset E which is a generalized Cantor set with a certain density property having Hausdorff dimension $\log(n+1)/\log(1+2c_0)$, where c_0

is the best constant in Markov's inequality on F for first degree polynomials which vanish at the center. Hence, from the algebraic characterization (2) of Markov's inequality for a set F we can immediately make a statement on the density and local Hausdorff dimension of F.

1.5. There is a partial converse of Theorem 2 with H^s changed to H^s_∞ as indicated in Remark 4.

THEOREM 3. *Let F be a closed, non-empty subset of $\mathbb{R}^n$. Assume that for some $s > n-1$ there exists a constant $c_2 > 0$ so that*

$$\frac{H^s_\infty(F \cap B(x_0, r))}{(2r)^s} \geqslant c_2,$$

for every $x_0 \in F$ and $r \in (0, 1]$. Then F preserves Markov's inequality.

Proof. The proof is closely related to the proof of Theorem 3 on p. 39 in [3]. Take $x_0 \in F$, $r \in (0, 1]$, $B = B(x_0, r)$ and a band like that in Proposition 4 in Section 1.1,

$$S := \{x \in \mathbb{R}^n \colon |b \cdot (x - x_0)| < \varepsilon r\},$$

where $b \in \mathbb{R}^n$, $|b| = 1$, and ε is a small positive number. We can cover $B \cap S$ by cubes with side εr so that we need at most $c_3 r^{n-1} \cdot \varepsilon r/(\varepsilon r)^n = c_3 \varepsilon^{1-n}$, $c_3 = c_3(n)$, cubes. Hence,

$$H^s_\infty(B \cap S) \leqslant c_3 \varepsilon^{1-n} \cdot (\varepsilon r \sqrt{n})^s = c_4 r^s \varepsilon^{s-(n-1)}, \qquad c_4 = c_4(n),$$

which tends to zero with ε since $s > n-1$. Combined with the assumptions on F this gives for any B and any S, if ε is small enough,

$$H^s_\infty(F \cap B) \geqslant c_2(2r)^s > c_4 r^s \varepsilon^{s-(n-1)} \geqslant H^s_\infty(B \cap S).$$

We choose such an ε and infer from the last chain of inequalities that $F \cap B$ is not a subset of $B \cap S$ for any B or S. Theorem 3 now follows from the geometric characterization in Proposition 4.

2. Sets Preserving a Generalized Markov Inequality

2.1. In this section we give an example of a cusp domain which does not satisfy Markov's inequality, and see that a weaker inequality, (7) below, in a natural way takes the place of Markov's inequality. For a constant $\lambda > 1$ let F be the cusp domain

$$F := \{(x, y) \in \mathbb{R}^2 \colon |y| \leqslant x^\lambda,\ 0 \leqslant x \leqslant 1\}. \qquad (6)$$

We can immediately see that F does not preserve Markov's inequality by checking with the geometric criterion in Proposition 4 at zero. Instead we have the following inequality which is closely related to the result in [1]:

PROPOSITION 8. *Let Q_r, $r>0$, be the square with side $2r$ and center at the origin, and let F be the cusp domain (6) where $\lambda>1$. Then, for every positive integer k there exists a constant $c(k)$ so that*

$$\|\nabla P\|_{Q_r \cap F} \leqslant \frac{c(k)}{r^\lambda} \|P\|_{Q_r \cap F}, \tag{7}$$

for all $P \in \mathscr{P}_k$ and $r \in (0, 1]$, and (7) is optimal as concerns the dependence on r.

Proof. (1) There are different ways to realize that there exists a constant $c(k)$ such that (7) holds for $r=1$ and all $P \in \mathscr{P}_k$; see for instance [1] or [2, Proposition 4], or prove it by using Markov's inequality followed by the formula immediately before Proposition 10. Now, for $r \in (0, 1]$, introduce the function $g: \mathbb{R}^2 \to \mathbb{R}^2$ defined by $(x, y) \mapsto (x/r, y/r^\lambda)$. For a given $P \in \mathscr{P}_k$ introduce $P_1 := P \circ g^{-1}$. Then $P = P_1 \circ g$ and by using (7) for $r=1$ we get

$$\left\|\frac{\partial P}{\partial 2}\right\|_{Q_r \cap F} = \frac{1}{r^\lambda} \left\|\frac{\partial P_1}{\partial 2}\right\|_{Q_1 \cap F} \leqslant \frac{c(k)}{r^\lambda} \|P_1\|_{Q_1 \cap F} = \frac{c(k)}{r^\lambda} \|P\|_{Q_r \cap F}.$$

Analogously we get

$$\left\|\frac{\partial P}{\partial 1}\right\|_{Q_r \cap F} \leqslant \frac{c(k)}{r} \|P\|_{Q_r \cap F}, \qquad P \in \mathscr{P}_k,\ r \in (0, 1], \tag{8}$$

and (7) is proved.

(2) By choosing $P(x, y) = y^k$ for any positive integer k we see that

$$\|P'_y\|_{Q_r \cap F} / \|P\|_{Q_r \cap F} = k/r^\lambda,$$

proving the optimality in r, and hence the proposition.

From [1] we see that $c(k) \leqslant 8\lambda e^4 k^{2\lambda}$ and that this constant is optimal as concerns the exponent 2λ in the expression $k^{2\lambda}$. From (8) we also see the difference between the size of P'_x and P'_y; the factor $1/r^\lambda$ in (7) is needed for P'_y but not for P'_x where $1/r$ is sufficient, a fact which is natural from the geometry of F.

By using (7) we deduce, in the following proposition, a condition on F of the type used in Proposition 2; we remark that the method of proof applies to more general sets F.

PROPOSITION 9. *Let Q_r, F, and λ be as in Proposition* 8. *Then, for every positive integer k, there exists a constant $c(k)$ such that*

$$\|P\|_{Q_r} \leqslant \frac{c(k)}{r^{k(\lambda-1)}} \|P\|_{Q_r \cap F}, \tag{9}$$

for all $P \in \mathscr{P}_k$ and $r \in (0, 1]$, and (9) *is optimal as concerns the dependence on r.*

Proof. (1) If $P(x) = \sum a_j x^j$, $|j| \leqslant k$, $x = (x_1, x_2) \in \mathbb{R}^2$, repeated application of (7) gives

$$|a_j| = \left|\frac{D^j P(0)}{j!}\right| \leqslant \frac{c_1(k)}{r^{\lambda|j|}} \|P\|_{Q_r \cap F},$$

and, consequently,

$$\|P\|_{Q_r} \leqslant \sum_{|j| \leqslant k} |a_j| \, r^{|j|} \leqslant \frac{c_2(k)}{r^{k(\lambda-1)}} \|P\|_{F \cap Q_r}.$$

(2) As in the previous proposition the optimality follows by considering $P(x, y) = y^k$, and the proof is complete.

We see that when λ tends to 1 the cusp in (6) vanishes and (7) turns into Markov's inequality on F. Proposition 8 is a reason to study sets preserving a generalized Markov inequality. Another such reason is given in [11] where sets satisfying an inequality related to (9) are studied with respect to unique polynomial interpolation and approximation.

2.2. Motivated by Section 2.1 we study the following condition on a closed, non-empty subset F of $\mathbb{R}^n$, where k is a positive integer and $\mu(k)$ a non-negative number: There exists a constant $c(k) = c(F, n, k)$ such that

$$\|P\|_B \leqslant \frac{c(k)}{r^{\mu(k)}} \|P\|_{B \cap F}, \tag{10}$$

for all $P \in \mathscr{P}_k$ and $B = B(x_0, r)$, $x_0 \in F$, $0 < r \leqslant 1$. In condition (10) the balls B may be changed to cubes Q and the condition $r \leqslant 1$ to $r \leqslant r_0$, where r_0 is a fixed positive number, without changing the class of sets F satisfying the condition. This follows by applying the inequality (see for instance [7, Lemma 1] or use the fact that all norms on the finite dimensional vector space $\mathscr{P}_k$ are equivalent)

$$\|P\|_{B(x_0, r)} \leqslant c(n, k, a) \, \|P\|_{B(x_0, ar)}, \qquad 0 < a < 1, \; P \in \mathscr{P}_k,$$

and the analogous inequality for cubes.

PROPOSITION 10. *F satisfies (10) for a certain k and $\mu(k)$, $\mu(k) \geqslant 0$, if and only if there exists a constant $c'(k) = c'(F, n, k)$ such that*

$$\|\nabla P\|_B \leqslant \frac{c'(k)}{r^{1+\mu(k)}} \|P\|_{B \cap F}, \tag{11}$$

for all $P \in \mathscr{P}_k$ and $B = B(x_0, r)$, $x_0 \in F$, $r \leqslant 1$.

Proof. (1) (11) $\Rightarrow$ (10). By the mean-value theorem we get, for $x \in B$,

$$|P(x)| \leqslant |P(x) - P(x_0)| + |P(x_0)| \leqslant r\|\nabla P\|_B + |P(x_0)|,$$

and an application of (11) gives (10).

(2) (10) $\Rightarrow$ (11). Markov's inequality

$$\|\nabla P\|_B \leqslant \frac{c_1(n, k)}{r} \|P\|_B$$

combined with (10) gives (11), and the proof is complete.

We remark that if (10) holds for $k = 1$ with $\mu(k) = \mu(1) = 0$, then F preserves Markov's inequality, by Proposition 3. In particular, this means that (10) holds for all positive integers k with $\mu(k) = 0$. When $\mu(1) > 0$ the typical situation is that $\mu(k)$ in (10) would be $k\mu(1)$ as illustrated for the cusp domain in Proposition 9. This means that we do not have an analogue of Proposition 3 for the generalized Markov inequality (11) and (10). On the other hand we do have a geometric characterization of (10) and (11) for $k = 1$ for instance along the lines of Proposition 7 (see (14) for $n = 1$).

2.3. We now turn to a more detailed study of (10) for the case $k = 1$. Let λ and c be numbers larger than 1 and let $\mathscr{F}(c, \lambda)$ be the class of non-empty subsets F of $\mathbb{R}^n$ such that, for all $B = B(x_0, r)$, $x_0 \in F$, $0 < r \leqslant 1$,

$$\|P\|_B \leqslant \frac{c}{r^{\lambda - 1}} \|P\|_{B \cap F}, \qquad \text{for all} \quad P \in \mathscr{P}_1. \tag{12}$$

Remark 6. The same class, $\mathscr{F}(c, \lambda)$, is defined if the inequality (12) is replaced by

$$\|\nabla P\|_{B \cap F} \leqslant \frac{c}{r^{\lambda}} \|P\|_{B \cap F}, \qquad P \in \mathscr{P}_1,$$

but with another constant c.

Remark 7. If $\lambda=1$ then the sets in $\mathscr{F}(c,\lambda)$ preserve Markov's inequality (see Section 1, Proposition 3).

PROPOSITION 11. *If $\lambda>1$ and $c>1$ then there exists a set in $\mathscr{F}(c,\lambda)$ of Hausdorff dimension* 0.

We prove Proposition 11 in the one dimensional case. However, the proof could be generalized to several dimensions.

Proof. We prove the existence of a set $F\subset\mathbb{R}^1$ in $\mathscr{F}(c,\lambda)$ which is of Hausdorff dimension 0 in four steps:

(1) We construct a set E as the closure of a countable union of successively chosen finite sets.

(2) We show that E has Hausdorff dimension 0.

(3) We prove that $E\in\mathscr{F}(3c,\mu)$, $\lambda>\mu>1$.

(4) We use the set E to construct a set F such that $F\in\mathscr{F}(c,\lambda)$ and $\dim_H(F)=0$.

In parallel with the construction of E we construct a continuous 1–1 mapping f of E onto the interval $[0,1]$. Through this parallel construction and by expressing numbers in $[0,1]$ in the binary system the proof will be easy to follow.

Step 1. We choose μ, $1<\mu<\lambda$, and introduce the notation

$$M(k)=\left(\frac{1}{c}\right)^{1+\mu+\mu^2+\cdots+\mu^k} \tag{13}$$

since this quantity $M(k)$ will appear often in the proof. Introduce

$$E_0=\{0\} \quad \text{and} \quad f\colon 0\mapsto 0.0, \quad \text{and,} \quad \text{for } k=0,1,2,...,$$
$$E_{k+1}=E_k\cup[E_k+M(k)] \quad \text{and,} \quad \text{for } a\in E_k,$$
$$f\colon a+M(k)\mapsto f(a)+0.0\cdots 01\ (=f(a)+2^{-k-1}).$$

Here $E_k+M(k)$ is E_k translated the distance $M(k)$. For instance, we see that E_1 consists of the points 0 and $1/c$ and that $f(1/c)=0.1$. Let E be the closure of $\bigcup E_k$. Because of the construction f will map, continuously and 1–1, a dense subset of E onto that dense subset of $[0,1]$ which consists of all numbers in $[0,1)$ with a finite number of ones in their binary expansion. The notation f will also be used for the unique extension of f to a 1–1 and continuous mapping of E onto $[0,1]$. The inverse of f is also continuous and 1–1 and we denote it by g.

Step 2. The construction and (13) shows that g is increasing and has two important properties:

(i) $0 \leqslant g(0.a_1 \cdots a_k a_{k+1} \cdots) - g(0.a_1 a_2 \cdots a_k 00 \cdots) \leqslant \sum_{i=k}^{\infty} M(i) < (c/(c-1)) \cdot M(k)$, and

(ii) $g(0.a_1 a_2 \cdots a_k 1 a_{k+2} \cdots) - g(0.a_1 a_2 \cdots a_k 0 a_{k+2} \cdots) = M(k)$.

From (i) it follows that E is covered by 2^k intervals of length

$$v(k) = \frac{cM(k)}{c-1}.$$

Hence, for every $\delta > 0$ and for every $v(k) \leqslant \delta$, we can estimate

$$H_\delta^s(E) \leqslant 2^k (v(k))^s,$$

which goes to zero for each positive s as k tends to infinity, since $\mu > 1$. Consequently $H_\delta^s(E) = 0$ for every $\delta > 0$, i.e., $H^s(E) = 0$ for $s > 0$, and hence the Hausdorff dimension of E is zero.

Step 3. If we can prove the following geometric property for the points in E then it follows that E belongs to $\mathscr{F}(3c, \mu)$. For each $x_0 \in E$ and $r \in (0, 1]$ there exists an $x \in E$ such that

$$\frac{r^\mu}{c} \leqslant |x - x_0| \leqslant r. \tag{14}$$

The reason this implies $E \in \mathscr{F}(3c, \mu)$ is as follows:

Take any non-constant $P \in \mathscr{P}_1$ and normalize so that $P'(x_0) = 1$. Then, for $B = B(x_0, r)$, we obtain by the geometric property

$$\|P\|_B = \max_B |P(x_0) + (x - x_0)| \leqslant |P(x_0)| + \frac{cr \max_{B \cap E} |x - x_0|}{r^\mu}.$$

By here writing $x - x_0$ as $P(x) - P(x_0)$, a trivial estimate shows that $E \in \mathscr{F}(3c, \mu)$.

We finish Step 3 by proving the desired geometric property of E. Assume $x_0 \in E$. Then there exists an $a \in [0, 1]$ such that

$$g(a) = x_0.$$

We now use that given $c > 1$, $r \in (0, 1/c]$, and $\mu > 1$ there exists a positive integer k such that

$$M(k) < r \leqslant M(k-1);$$

we put $k=0$ if $1/c<r\leqslant 1$. This gives

$$r^{\mu}/c\leqslant M(k)<r \tag{15}$$

and by property (ii) for g there exists for the given $g(a)\in E$ a $g(b)\in E$ such that

$$|g(a)-g(b)|=M(k). \tag{16}$$

Put $x=g(b)$ and we get, using (15) and (16), that the geometric property (14) holds.

Step 4. Let $c>1$ and $\lambda>1$ be given. Choose μ, $1<\mu<\lambda$, and let $r_0=(1/3)^{1/(\lambda-\mu)}$. Use E from Step 1 to form, for a large positive integer m,

$$F_i=E+\frac{i}{m},\qquad i \text{ integer},\qquad \text{and}\qquad F(m)=\bigcup_{i=-\infty}^{\infty}F_i.$$

Since $\dim_H(E)=0$ it follows that $\dim_H(F(m))=0$. We now prove that there exists an m such that $F(m)\in\mathscr{F}(c,\lambda)$, and then by putting $F=F(m)$ Step 4 will be proved. It is easy to see that by choosing m large enough, $m\geqslant m_0$, we can make

$$\inf\left\{\frac{\|P\|_{B\cap F(m)}}{\|P\|_B}:P\in\mathscr{P}_1,\ P\neq 0\right\}$$

larger than any prescribed number less than 1, independently of $B=B(x_0,r)$ if $x_0\in F(m)$ and $r\geqslant r_0$.

It follows that (12) holds for $F=F(m_0)$ for those balls $B=B(x_0,r)$, where $x_0\in F(m_0)$ and $r\geqslant r_0$. It remains to prove that (12) holds also for $r<r_0$. But we know from Step 3 that $E\in\mathscr{F}(3c,\mu)$ was achieved using a geometric property for the points in E, and since $F(m_0)$ by construction has the same property we conclude that $F(m_0)\in\mathscr{F}(3c,\mu)$. If we use this, and the fact that we estimate only for $r<r_0$, we get

$$\|P\|_B\leqslant\frac{3c}{r^{\mu-1}}\|P\|_{F(m_0)\cap B}\leqslant\frac{c}{r^{\lambda-1}}\|P\|_{F(m_0)\cap B}$$

which is (12) for $r<r_0$. By that Step 4 and the proposition is proved.

Remark 8. An inspection of the proof of Step 2 shows that we can, in fact, say more on the size of the set constructed. Given that $d>\log 2/\log\lambda$, $\lambda>1$, we can construct a set in $\mathscr{F}(c,\lambda)$, $c>1$, having Hausdorff measure zero with respect to $h(x)=(\log 1/x)^{-d}$.

References

1. P. Goetgheluck, Inégalité de Markov dans les ensembles effilés, *J. Approx. Theory* **30** (1980), 149–154.
2. A. Jonsson, Markov's inequality and local polynomial approximation, *in* "Lecture Notes in Mathematics," Vol. 1302, pp. 303–316, Springer-Verlag, Berlin, 1986.
3. A. Jonsson and H. Wallin, "Function Spaces on Subsets of $\mathbb{R}^n$," Mathematical Reports, Vol. 2, Part 1, Harwood Academic, London, 1984.
4. M. A. Martin and P. Mattila, k-Dimensional regularity classifications, *Trans. Amer. Math. Soc.* **305** (1988), 293–315.
5. P. A. P. Moran, Additive functions of intervals and Hausdorff measure. *Proc. Cambridge Philos. Soc.* **42** (1946), 15–23.
6. W. Pleśniak, Markov's inequality and the existence of an extension operator for C^∞ functions, *J. Approx. Theory* **61** (1990), 106–117.
7. H. Wallin, "Self Similarity, Markov's Inequality and d-Sets," No. 7, Department of Mathematics, University of Umeå, 1987.
8. H. Wallin, Interpolating and orthogonal polynomials on fractals, *Constr. Approx.* **5** (1989), 137–150.
9. P. Wingren, Lipschitz spaces and interpolating polynomials on subsets of Euclidean space, *in* "Lecture Notes in Mathematics," Vol. 1302, pp. 424–435, Springer-Verlag, Berlin, 1986.
10. P. Wingren, Simultaneous multivariate polynomial interpolation and approximation on fractal sets, *in* "Progress in Approximation Theory," pp. 843–856, Academic Press, San Diego, 1991.
11. P. Wingren, "A Note on Polynomial Interpolation," No. 2, Department of Mathematics, University of Umeå, 1990.
12. A. P. Wójcik, Some remarks on Bernstein's theorems, *J. Approx. Theory* **67** (1991), 252–269.