Density and covering properties of intervals of $\mathbf{R}^n$

Tamás Keleti

February 11, 2001

Abstract

The key result of this paper is the existence of functions $\rho_n(h)$ for which whenever H is a (Lebesgue) measurable subset of the n-dimensional unit cube I^n with measure $|H| > h$ and $\mathcal{R}$ is a class of subintervals (n-dimensional axis-parallel rectangles) of I^n that covers H, then there exists an interval $R \in \mathcal{R}$ in which the density of H is greater than $\rho_n(h)$; that is, $\frac{|H \cap R|}{|R|} > \rho_n(h)$ $\left(= \left(\frac{h}{2n}\right)^n \right)$. We show how we can use this result for finding 4 points of a measurable subset of the unit square such that they are the vertices of an axis-parallel rectangle that has quite large intersection with the original set. We introduce and investigate density and covering properties of classes of subsets of $\mathbf{R}^n$. As a consequence we get a covering property of the class of intervals of $\mathbf{R}^n$: if $\mathcal{R}$ is a family of n-dimensional intervals with $|\cup \mathcal{R}| < \infty$ then there is a finite sequence $R_1, \ldots, R_m \in \mathcal{R}$ such that $|\cup_{k=1}^m R_k| \geq (1-\varepsilon)|\cup \mathcal{R}|$ and $\|\sum_{k=1}^m \chi_{R_k}\|_q \leq C(n,q,\varepsilon)|\cup \mathcal{R}|^{1/q}$.

1 Introduction

While the author was working on a modified problem of A. Carbery, the following question arose:

If a measurable subset of the unit square is covered by axis-parallel rectangles (contained in the unit square) such that its density is small in each rectangle, can we conclude that the set itself must have small measure?

First note that if we allow any (not necessary axis-parallel) rectangles then the answer is negative. Indeed, a closed subset of a Nikodym set (a set in the unit square with measure one such that for each point of the set there is a straight line intersecting the set only in that single point, see e. g. [6]) with measure $1-\varepsilon$ can be easily covered by rectangles such that the density of the subset is less then ε in each rectangle.

1991 *Mathematical Reviews Classification.* Primary 28A75; Secondary 42B25.

However, as we shall prove (Theorem 2.1), for axis-parallel rectangles (even in n-dimension) the answer is affirmative. Then we can easily get a similar result for sets not necessarily in the unit cube (Theorem 2.5). Using this property of the intervals (the n-dimensional axis-parallel rectangles) of $\mathbf{R}^n$ - that we shall call the "minimal density property" or shortly MDP - we can also prove a covering property of the intervals of $\mathbf{R}^n$ (Theorem 2.6).

In Section 3 we present a result about the modified problem of A. Carbery (asked by I. Gyöngy) that motivated our investigation. Namely, we use the minimal density property of the axis-parallel rectangles to find an (axis-parallel) rectangle with vertices in a given set with large intersection with this set. For this we will also need a result for the original problem of A. Carbery.

In Section 4 we investigate the classes of subsets of $\mathbf{R}^n$ that have the minimal density property. We investigate how this property relates to some (old and new) covering properties. We show how we can improve some covering properties (the V_q property) for classes satisfying the MDP. Thus we can prove a strong covering property of the intervals of R^n.

Most of the measure theoretic results can be equivalently formulated as combinatorial ones, in the sense that the measurable sets and the intervals may be assumed to be finite unions of dyadic cubes and the coverings may be assumed to be finite. Nevertheless, the proof of our key result (Theorem 2.1) uses methods of analysis. We investigate a minimal operator analogue to the well known Hardy-Littlewood maximal operator (see e.g. [6] or [7]).

Notation 1.1 We denote by I^n the n-dimensional open unit cube; that is, $I^n = (0,1)\times\ldots\times(0,1)$. By an *interval* of R^n we mean an n-dimensional axis-parallel open rectangle: the Cartesian product of n open (1-dimensional) intervals. We denote by $\mathcal{I}^n$ the class of intervals of $\mathbf{R}^n$ and by $\mathcal{I}_0^n$ the class of (n-dimensional) subintervals of I^n.

We denote the (Lebesgue) measure and the closure of a set $A \subset \mathbf{R}^n$ by $|A|$ and $\overline{A}$, resp. By the density of A in B (with $|B|>0$) we mean $\frac{|A\cap B|}{|B|}$.

2 Intervals of $\mathbf{R}^n$

Theorem 2.1 *If H is a measurable subset of I^n with $|H|>h$ and $\mathcal{R}$ is a class of intervals in I^n that covers H, then there exists an interval $R\in\mathcal{R}$ in which the density of H is greater than $(\frac{h}{2n})^n$; that is,*

$$\frac{|H\cap R|}{|R|} > \left(\frac{h}{2n}\right)^n.$$

Proof. We prove the statement by induction on n. Let $n = 1$. Take a finite subclass of $\mathcal{R}$ that intersects H in a set of measure greater than h. It is well known that from a finite class of intervals one can always select two subclasses of disjoint intervals such that the union of the selected intervals is the same as the union of the whole class. Then, in our case, at least one of the selected classes of disjoint intervals intersects H in a set of measure greater than $h/2$. Thus in at least one of these intervals the density of H must be greater than $h/2$.

Assume that the statement is true for $n-1$. Since we can find a closed set $H' \subset H$ with $|H'| > h$ we can assume that H is closed. Then we can cover every point of $I^n \setminus H$ by an interval disjoint to H, thus we can assume that $\mathcal{R}$ covers the whole I^n.

Let

$$m(x_1, \ldots, x_n) = \inf \left\{ \frac{|H \cap (x_1 \times T)|}{|T|} \; : \; T \in \mathcal{I}_0^{n-1}, \; (x_2, \ldots, x_n) \in T \right\}.$$

Standard arguments show (see e.g. [5]) the measurability of the function $m : I^n \to [0, 1]$.

Suppose that the density of H is at most b in every $R \in \mathcal{R}$. Then we prove that

$$2b \geq \int_{I^n} m > 2\left(\frac{h}{2n}\right)^n,$$

which clearly implies our statement.

- $2b \geq \int_{I^n} m$

  Fix $x_2, \ldots, x_n \in I$. For a $t \in I$ let $K_t \times T_t$ $(K_t \in \mathcal{I}_0^1, T_t \in \mathcal{I}_0^{n-1})$ be an interval in $\mathcal{R}$ that covers $(t, x_2, \ldots, x_n)$. By definition,

  $$m(s, x_2, \ldots, x_n) \leq \frac{|H \cap (s \times T_t)|}{|T_t|} \qquad \text{(for any } s \in I\text{)}.$$

  Thus, integrating and using that the density of H in $K_t \times T_t$ is at most b, we get

  $$\int_{K_t} m(s, x_2, \ldots, x_n) ds \leq |K_t| \frac{|H \cap (K_t \times T_t)|}{|K_t \times T_t|} \leq |K_t| b. \tag{1}$$

  The intervals K_t $(t \in I)$ cover I, so, taking a finite class of intervals K_t that covers I except a set of measure at most ε and selecting two subclasses of disjoint intervals with the same union, we get intervals

$K_{t_1}, \ldots, K_{t_m}$ that covers I except a set of measure at most ε such that every point is covered at most twice. From (1) we get

$$\sum_{i=1}^{m} \int_{K_{t_i}} m(s, x_2, \ldots, x_n) ds \leq \sum_{i=1}^{m} |K_{t_i}| b. \tag{2}$$

Since the intervals $K_{t_1}, \ldots, K_{t_m}$ cover I except a set of measure at most ε and $0 \leq m \leq 1$ we get that the left-hand side of (2) is at least $\int_I m(s, x_2, \ldots, x_n) ds - \varepsilon$. On the other hand, every point is covered at most twice, so the right-hand side is at most $2b$. Therefore we have $\int_I m(s, x_2, \ldots, x_n) ds \leq 2b$, which, integrated with respect to $x_2, \ldots, x_n$, gives the inequality we wanted to prove.

- $\int_{I^n} m > 2\left(\frac{h}{2n}\right)^n$

  Let

  $$A = \{x \in H : m(x) < a\} \quad \text{where} \quad a = \left(\frac{|H|}{2n}\right)^{n-1} = \left(\frac{\frac{n-1}{n}|H|}{2(n-1)}\right)^{n-1}.$$

  Fix $x_1 \in I$. Let $A^{x_1} = \{(x_2, \ldots, x_n) : (x_1, \ldots, x_n) \in A\}$. By definition, any $(x_2, \ldots, x_n) \in A^{x_1}$ is covered by a $T \in \mathcal{I}_0^{n-1}$ with

  $$\frac{|H \cap (x_1 \times T)|}{|T|} < a.$$

  Since $A \subset H$ it implies that A^{x_1} is covered by ($n-1$-dimensional) intervals in which its density is less than a. By our induction assumption, this implies that $|A^{x_1}| \leq \frac{n-1}{n}|H|$. Thus $|A| \leq \frac{n-1}{n}|H|$, which implies that $|H \setminus A| \geq |H|/n$.

  Using this, we get that

  $$\int_{I^n} m \geq \int_{H \setminus A} m \geq \int_{H \setminus A} a = |H \setminus A| a \geq \frac{|H|}{n}\left(\frac{|H|}{2n}\right)^{n-1} > 2\left(\frac{h}{2n}\right)^n,$$

  which completes the proof. □

Remark 2.2 Similar covering properties of intervals of the real line have been studied for very long time. The $n = 1$ case of Theorem 2.1 also follows from Youngs' First Covering Lemma ([10], 2. Lemma) from 1910, which says that if each point of a compact subset of the real line is the left-hand end-point of at least one interval then we can find a finite number of these intervals, non-overlapping, such that the measure of the non-covered part of the closed set is smaller than any fixed positive number.

Remark 2.3 The method we used in this proof is similar to the method used in [5] but, instead of Hardy-Littlewood maximal operator, we used the corresponding minimal operator. In fact, in the same way as in the proof of Theorem 2.1, we can also get a (very weak-type) inequality for the minimal operator. Namely, denoting the minimal operator associated to $\mathcal{I}_0^n$ by m_n (that is, $m_n f(x) = \inf\left\{\frac{1}{|R|}\int_R |f| : x \in R \in \mathcal{I}_0^n\right\}$ for any $f \in L_1(I^n)$), we can prove that $\rho_n\left(\int_{\{m_n f<b\}} f\right) \le b$ for any $f : I^n \to I$ measurable function and $b > 0$. From this we can easily obtain that

$$\int_{\{m_n f<b\}} |f| \le 2n\|f\|_\infty^{1-\frac{1}{n}} b^{\frac{1}{n}}$$

for any $f \in L_\infty(I^n)$ and $b > 0$.

A similar notion of minimal operator was introduced in [3].

Notation 2.4 We shall call the function $(\frac{h}{2n})^n$ (which appeared in Theorem 2.1) $\rho_n(h)$.

Theorem 2.5 *Suppose that H is a measurable subset of $\mathbf{R}^n$ with finite measure, $\mathcal{R}$ is a class of intervals of $\mathbf{R}^n$ that covers H and the density of H in $\cup\mathcal{R}$ is greater than $h > 0$. Then there exists an interval $R \in \mathcal{R}$ in which the density of H is greater than $\rho_n(h)$; that is,*

$$\frac{|H \cap R|}{|R|} > \rho_n(h) = \left(\frac{h}{2n}\right)^n.$$

Proof. We can assume that $\mathcal{R}$ is finite since one can select a finite subclass $\mathcal{R}' \subset \mathcal{R}$ such that the density of $H \cap (\cup\mathcal{R}')$ in $\cup\mathcal{R}'$ is still greater than h.

It is known (see e.g. [6] p. 70) that if G is an open bounded subset of $\mathbf{R}^n$ and K is a compact set with positive measure then there is a disjoint sequence $\{K_k\}$ of sets homothetic to K contained in G such that $|G \setminus \cup_{k=1}^\infty K_k| = 0$.

Applying this for $G = I^n$ and $K = \overline{\cup\mathcal{R}}$, we get the sequence $\{K_k\}$ and homothecies $\phi_k : K \to K_k$. Let $H_k = \phi_k(H)$, $\mathcal{R}_k = \phi_k(\mathcal{R})$. Then $\mathcal{R}^* = \cup_{k=1}^\infty \mathcal{R}_k$ covers $H^* = \cup_{k=1}^\infty H_k$. Clearly $|H^*| = \sum |H_k|$, $1 = \sum |K_k|$ and $|H_k|/|K_k| = |H|/|K|$, so $|H^*| = |H|/|K| > h$.

Applying Theorem 2.1, we can select an $R' \in \cup_{k=1}^\infty \mathcal{R}_k$ in which the density of H^* is greater than $\rho_n(h)$. If $R' \in \mathcal{R}_k$ then the density of H in $R = \phi_k^{-1}(R') \in \mathcal{R}$ is also greater than $\rho_n(h)$. □

Theorem 2.6 *For each $n \in \mathbf{N}$ there is a function $C_n : \mathbf{R}^+ \to \mathbf{R}^+$ such that for any $\mathcal{R} \subset \mathcal{I}^n$ with $|\cup \mathcal{R}| < \infty$ and $\varepsilon > 0$ there exist $R_1, \ldots, R_m \in \mathcal{R}$ for which*

$$(i) \qquad \frac{|\cup \mathcal{R} \setminus \cup_{k=1}^m R_k|}{|\cup \mathcal{R}|} < \varepsilon$$

and

$$(ii) \qquad \frac{\sum_{k=1}^m |R_k|}{|\cup \mathcal{R}|} < C_n(\varepsilon).$$

Proof. Let

$$t = |\cup \mathcal{R}|, \qquad \mathcal{R}_\delta = \{R \in \mathcal{R} : |R| > \delta\} \quad \text{and} \quad H^\delta = \cup \mathcal{R}_\delta.$$

Then $\cup\mathcal{R} = \cup_{\delta>0} H^\delta$, so we can choose $\delta > 0$ such that $|\cup \mathcal{R} \setminus H^\delta| < \varepsilon t/2$.

Let $H_1 = H^\delta$. Assume that $k \geq 1$ and $H_k \subset H^\delta$ is already defined. Let $a_k = |H_k|/(2t)$.

If $a_k < \varepsilon/4$ then let $m = k - 1$ and the procedure is finished.

Otherwise, applying Theorem 2.5 for the $\mathcal{R}^\delta$ covering of H_k, we get $R_k \in \mathcal{R}^\delta$ with

$$\frac{|R_k \cap H_k|}{|R_k|} > \rho_n(a_k).$$

Let $H_{k+1} = H_k \setminus R_k$.

We claim that this procedure finishes after a finite number of steps. Indeed, if $a_k \geq \varepsilon/4$ for every k then

$$|R_k \cap H_k| > \rho_n(\varepsilon/4)|R_k| > \rho_n(\varepsilon/4)\delta,$$

which is impossible since the sets $R_k \cap H_k$ are disjoint subsets of a set with finite measure.

Since $H_{m+1} = H^\delta \setminus \cup_{k=1}^m R_k$, $a_{m+1} = |H_{m+1}|/(2t) < \varepsilon/4$ and $|\cup\mathcal{R} \setminus H^\delta| < \varepsilon t/2$ we get

$$\frac{|\cup \mathcal{R} \setminus \cup_{k=1}^m R_k|}{t} < \varepsilon,$$

which means that (i) is satisfied.

Let

$$d_k = \frac{|R_k \cap H_k|}{2t} \quad (k = 1, \ldots, m).$$

Then $d_k = a_k - a_{k+1}$ and $\frac{|R_k|}{2t} < \frac{d_k}{\rho_n(a_k)}$. Thus

$$\frac{\sum_{k=1}^m |R_k|}{2t} < \sum_{k=1}^m \frac{d_k}{\rho_n(a_k)} = \sum_{k=1}^m (a_k - a_{k+1}) \frac{1}{\rho_n(a_k)}.$$

Since $\rho_n(x)$ is increasing and $a_k \geq \varepsilon/4$ the right-hand side is a lower estimate of the integral of the function $\min(1/\rho_n(x), 1/\rho_n(\varepsilon/4))$ in the interval $[a_{m+1}, a_1]$, so we get

$$\frac{\sum_{k=1}^m |R_k|}{2t} < \int_{a_{m+1}}^{a_1} \min\left(\frac{1}{\rho_n(x)}, \frac{1}{\rho_n(\frac{\varepsilon}{4})}\right) dx < \frac{\varepsilon/4}{\rho_n(\frac{\varepsilon}{4})} + \int_{\varepsilon/4}^{1/2} \frac{1}{\rho_n(x)} dx = \frac{C_n(\varepsilon)}{2}.$$

Therefore (ii) is also satisfied. □

Remark 2.7 Using that $\rho_n(x) = (x/(2n))^n$, the proof above gives $C_n(\varepsilon) < \frac{n}{4(n-1)}(8n)^n(1/\varepsilon)^{n-1}$ for $n \geq 2$. For $n = 1$, using the same argument as in the very first part of the proof of Theorem 2.1, we get $C_1 = 2$.

Example 2.8 Let $0 < \delta < 1$ and let $\mathcal{R}$ be the class of axis-parallel unit squares with lower-left vertices on the segment $\{(x, y) : x + y = 0, -1 \leq x \leq 0\}$. Let $H = \{(x, y) : 0 \leq x + y \leq \delta\} \cap \cup\mathcal{R}$. Then $\mathcal{R}$ covers H and we have $|H| > \delta$ and $|\cup\mathcal{R}| = 3$. Thus the density of H in $\cup\mathcal{R}$ is greater than $\delta/3$ but its density is $\delta^2/2 = 4.5(\delta/3)^2$ in any $R \in \mathcal{R}$. Therefore Theorem 2.5 (and consequently Theorem 2.1) cannot be true with $\rho_2'(h) = 4.5h^2$, which means that only the constant can be improved in these results (for $n = 2$). (Slightly modifying this construction, we can also prove that no function greater than $h^2/2$ can be good either.)

One can check that if $R_1, \ldots, R_m \in \mathcal{R}$ (where $\mathcal{R}$ is the same as above) then $|\cup\mathcal{R} \setminus \cup_{k=1}^m R_k| > 1/m$. Thus, whenever $|\cup\mathcal{R} \setminus \cup_{k=1}^m R_k|/|\cup\mathcal{R}| < \varepsilon$, we have $m > \frac{1}{3\varepsilon}$, hence $(\sum_{k=1}^m |R_k|)/|\cup\mathcal{R}| > \frac{1}{9\varepsilon}$. Therefore Theorem 2.6 would not be true for $C_2(\varepsilon) \leq \frac{1}{9\varepsilon}$. On the other hand, according to Remark 2.7, Theorem 2.6 holds for $C_2(\varepsilon) = \frac{128}{\varepsilon}$.

In a similar way (taking n-dimensional axis-parallel cubes with "lower-left" vertices on a not too small domain of the hyperplane $\{x_1 + \ldots + x_n = 0\}$) we can show that, in higher dimensions as well, we have the best possible exponents in the above mentioned results.

3 Application

Recently A. Carbery asked the following question:

For which functions $a : [0, 1] \to [0, 1]$ *is it true that*

(∗) *if H is a measurable subset of I^2 then one can always find 4 points of H such that they are the vertices of a (2-dimensional) interval with area at least $a(|H|)$?*

This question led I. Gyöngy to ask the following question:
For which functions $f : [0,1] \to [0,1]$ is it true that

(∗∗) *if H is a measurable subset of I^2 then one can always find 4 points of H such that they are the vertices of a (2-dimensional) interval R such that $|R \cap H| \geq f(|H|)$?*

Clearly it is more difficult to satisfy (∗∗) then (∗). However, we shall see that, using Theorem 2.1, it is easy to obtain a function satisfying (∗∗) from a function that satisfies (∗).

A. Carbery, M. Christ and J. Wright [1] proved that $a(h) = ch^2/\log(1/h)$ (for a suitable $c > 0$ and h small enough) satisfies (∗). For the sake of completeness, we sketch a proof of this result.

Proposition 3.1 *If a measurable set $H \subset I^2$ with measure u does not contain the 4 vertices of any interval with area at least v then we have*

$$u^2 \leq 2v \log \frac{1}{v} + v^2. \tag{3}$$

Proof. Since there exists a closed subset of H with measure arbitrarily close to u we can assume that H is closed. Let H_m be the union of those closed squares of the regular $m \times m$ subdivision of I^2 that intersect H. Only finitely many H_m can contain the 4 vertices of an interval with area at least v since otherwise, taking a subsequence in which all the 4 vertices converge, we would get an interval with area at least v and with vertices belonging to H. Let K_m be the set that we get by magnifying H_m with ratio m. Thus (for a fixed large m) K_m consists of at least $m^2 u$ (unit) squares and they form no (axis-parallel) rectangle with area at least $m^2 v$.

Let k_i be the number of squares (of K_m) in the i-th row. Let P be the number of the horizontal square pairs; that is,

$$P = \sum_{i=1}^{m} \binom{k_i}{2} \geq \frac{1}{2} \frac{\left(\sum_{i=1}^{m} k_i\right)^2}{m} - \frac{1}{2} \sum_{i=1}^{m} k_i \geq \frac{1}{2} m^3 (u^2 - o(1)). \tag{4}$$

The j-th and the $j+i$-th squares cannot be both in K_m in more than $m^2 v / i$ rows since otherwise K_m contains the vertices of a rectangle with area at least m^2v. Thus

$$\begin{aligned} P &\le \sum_{i=1}^{m-1}\sum_{j=1}^{m-i}\min(m, m^2v/i) = \sum_{i=1}^{[mv]}(m-i)m + \sum_{i=[mv]+1}^{m}(m-i)m^2v/i \\ &= o(m^3) + mv(m-(mv/2))m + m^3v\log(1/v) - (m-mv)m^2v \\ &= m^3(v\log(1/v) + v^2/2 + o(1)) \qquad (5) \end{aligned}$$

Combining (4) and (5) and letting m tend to infinity, we get the inequality (3). □

Corollary 3.2 *The function $a(u) = cu^2/\log(1/u)$ (if $u \le \delta < 1$ and $a(u) = a(\delta)$ if $u > \delta$) satisfies $(*)$, where $c > 0$ depends on δ (and $c \to 1/4$ as $\delta \to 0$).* □

Example 3.3 Let H_m be the union of the diagonal squares of the regular $m \times m$ subdivision of the unit square. Then clearly $|H_m| = 1/m$ and each axis-parallel rectangle with vertices in H_m has area at most $\frac{1}{m^2}$. Thus a function that satisfies $(*)$ cannot be greater than u^2. It is unknown weather $a(u) = cu^2$ satisfies $(*)$ (for a sufficiently small $c > 0$).

Proposition 3.4 *If the function a satisfies $(*)$ then $f(h) = \rho_2(h/2)a(h/2)$ satisfies $(**)$, (where $\rho_2(h) = h^2/16$ is the function that appeared in Theorem 2.1 for $n = 2$).*

Proof. Let H be a measurable subset of I^2 with measure h and let

$$\mathcal{R} = \left\{\overline{R} : R \in \mathcal{I}_0^2 : |R| \ge a\left(\frac{h}{2}\right) \text{ and } \frac{|R\cap H|}{|R|} < \rho_2\left(\frac{h}{2}\right)\right\}.$$

Let $H' = H \cap \cup\mathcal{R}$.

The class $\mathcal{R}$ covers H' but the density of H is less than $\rho_2(h/2)$ in any $R \in \mathcal{R}$. Thus, by Theorem 2.1, we must have $|H'| \le h/2$. Hence $|H \setminus H'| \ge h/2$, so, using that the function a satisfies $(*)$, we can find an interval R with vertices in $H \setminus H'$ such that $|R| \ge a(h/2)$. Since $\overline{R}$ is not in $\mathcal{R}$ we get that $|R\cap H|/|R| \ge \rho_2(h/2)$. Thus $|R\cap H| \ge \rho_2(h/2)|R| \ge \rho_2(h/2)a(h/2) = f(h)$. □

Corollary 3.5 *The function $f(h) = c'h^4/\log(1/h)$ (if $h \le \delta < 1$ and $f(h) = f(\delta)$ if $h > \delta$) satisfies $(**)$, where c' depends only on δ.*

Proof. It follows from Corollary 3.2 and Proposition 3.4. □

Example 3.6 We use a construction of I. Reiman [9]. Let p be a prime number (or a power of a prime) and let $a_1, \dots, a_m$ be the points and $b_1, \dots, b_m$ be the lines of the (finite) projective plane of order p, where $m = p^2 + p + 1$. We take H_p to be the union of open squares of the regular $m \times m$ subdivision of the unit square as follows: we take the square in the i-th row and j-th column if and only if a_i is on b_j. Then, using that two lines meet only in one point, we have that whenever 4 points of H_p are the vertices of an axis-parallel rectangle R then the vertices must be in one row or in one column of the subdivision. On the other hand, each line contains $p+1$ points and each point is on $p+1$ lines, so we get that $h = |H_p| = (p+1)/m > m^{-1/2}$ and $|H \cap R| \leq (p+1)/m^2 \leq m^{-3/2} + m^{-2} < h^3 + h^4$. Therefore $(**)$ does not hold for the function $h^3 + h^4$ ($\sim h^3$).

Therefore the best exponent (or the infimum of the exponents) for functions satisfying $(**)$ is in the interval [3,4]. This is the best we currently know.

Remark 3.7 All positive results of this section can be easily generalized to n-dimensional spaces: Instead of Proposition 3.1, with a similar counting argument, we can prove by induction that if $H \subset I^n, |H| = u_n$ and H does not contain the 2^n vertices of any n-dimensional interval then we have $u_n \leq o\left(v_n^{\frac{1}{2^{n-1}+\alpha}}\right)$ (as $v_n \to 0$) for any $\alpha > 0$. Then we get that $a_n(u) = c_n u^{2^{n-1}+\alpha}$ satisfies the n-dimensional version of $(*)$ (for proper $c_n > 0$ depending only on n and α).

The proof of Proposition 3.4 clearly works in any dimension, hence the statement holds also in n-dimensions. Thus $f_n(h) = c_n' h^{n+2^{n-1}+\alpha}$ satisfies the n-dimensional version of $(**)$.

However, it is considerably more difficult to construct examples showing that we cannot have much better results than the above mentioned. The natural n-dimensional generalization of Example 3.3 (e. g. the union of those cubes of the regular $m \times \ldots \times m$ subdivision of the unit cube for which the sum of the coordinates is divisible by m) shows only that a function satisfying the n-dimensional version of $(*)$ cannot be greater than u^n. No natural generalization of Example 3.6 seems to be known.

By standard probabilistic method, it is easy to prove the following combinatorial result:

One can select $O(m^{n-n/2^{n-1}})$ points of the regular n-dimensional $m \times \ldots \times m$ lattice such that no 2^n of them are the vertices of an n-dimensional interval. Moreover, we can assume that we chose $O(m^{n-1-n/2^{n-1}})$ points of each $n-1$-dimensional $m \times \ldots \times m$ sublattice.

Then, taking the union of the corresponding open cubes of a regular subdivision of the unit cube, we get a set H with measure $O(1/m^{n/2^{n-1}})$ such that if the vertices of an n-dimensional interval R are in H then $|R| < 1/m$ and $|R \cap H| < O(1/m^{1+n/2^{n-1}})$. Thus we get $O(u^{2^{n-1}/n})$ and $O(u^{(2^{n-1}/n)+1})$ functions that do not satisfy the n-dimensional versions of $(*)$ and $(**)$, respectively; which are still quite far from our positive results.

One possible way to obtain better examples is to show that, as Erdős [4] conjectured, one can also select $O(m^{n-1/2^{n-1}})$ points of the regular n-dimensional $m \times \ldots \times m$ lattice such that no 2^n of them are the vertices of an n-dimensional interval.

Then we would have $O(u^{2^{n-1}})$ and $O(u^{2^{n-1}+1})$ functions that do not satisfy the n-dimensional versions of $(*)$ and $(**)$, respectively, which would be quite close to our positive results.

4 The minimal density property

Notation 4.1 We denote the L_q norm of a function $f : \mathbf{R}^n \to \mathbf{R}$ by $||f||_q$; that is, $||f||_q = (\int_{\mathbf{R}^n} |f|^q)^{1/q}$. The characteristic function of a set $A \subset \mathbf{R}^n$ is denoted by χ_A.

Definition 4.2 Let $\mathcal{B}$ be a class of nonempty open bounded subsets of $\mathbf{R}^n$ and $1 \le q \le \infty$.

- We say that $\mathcal{B}$ has the *minimal density property* (MDP) if there exists a function $\rho : \mathbf{R}^+ \to \mathbf{R}^+$ such that if $H \subset \mathbf{R}^n$ is measurable with finite measure, $\mathcal{R} \subset \mathcal{B}$ covers H and the density of H in $\cup\mathcal{R}$ is $d > 0$ then one can find an $R \in \mathcal{R}$ in which the density of H is greater than $\rho(d)$; that is,
$$\frac{|R \cap H|}{|R|} > \rho\left(\frac{|H|}{|\cup \mathcal{R}|}\right).$$

- The class $\mathcal{B}$ is said to have the *covering property* V_q (see [2]) if there exist constants $C < \infty$ and $c > 0$ such that for any $\mathcal{R} \subset \mathcal{B}$ with $|\cup \mathcal{R}| < \infty$ we can find $R_1, \ldots, R_m \in \mathcal{R}$ such that
$$\text{(i')} \quad |\cup_{k=1}^m R_k| \ge c\, |\cup \mathcal{R}| \quad \text{and} \quad \text{(ii)} \quad ||\sum_{k=1}^m \chi_{R_k}||_q \le C\, |\cup \mathcal{R}|^{1/q}.$$

- We say that $\mathcal{B}$ has the *complete covering property* V_q (CV$_q$) if there exists a function $C : \mathbf{R}^+ \to \mathbf{R}^+$ such that for any $\varepsilon > 0$ and $\mathcal{R} \subset \mathcal{B}$

with $|\cup\mathcal{R}| < \infty$ we can find $R_1, \ldots, R_m \in \mathcal{R}$ such that

$$\text{(i)} \quad |\cup_{k=1}^m R_k| \geq (1-\varepsilon)|\cup\mathcal{R}| \quad \text{and} \quad \text{(ii)} \quad \|\sum_{k=1}^m \chi_{R_k}\|_q \leq C(\varepsilon)|\cup\mathcal{R}|^{1/q}.$$

Remark 4.3 If $\mathcal{R}$ is a class of open sets with $|\cup\mathcal{R}| < \infty$ then for any $\varepsilon_1 > 0$ there exists a finite subclass $\mathcal{R}' \subset \mathcal{R}$ such that $|\cup\mathcal{R}'| \geq (1-\varepsilon_1)|\cup\mathcal{R}|$. (Indeed, since $\mathbf{R}^n$ is hereditary Lindelöf, there exist $R_1, R_2, \ldots \in \mathcal{R}$ such that $\cup_i R_i = \cup\mathcal{R}$, hence $\lim_{N\to\infty} |\cup_{j=1}^N R_j| = |\cup_i R_i| = |\cup\mathcal{R}|$.)

Therefore, if we want to prove any of the above mentioned properties, we can assume that $\mathcal{R}$ is finite.

Remark 4.4 Note that Theorem 2.5 and Theorem 2.6 state that $\mathcal{I}^n$ (the intervals of $\mathbf{R}^n$) has the minimal density property and the CV_1 property. In the proof of Theorem 2.6 we used only the minimal density property of $\mathcal{I}^n$, so we proved that MDP implies CV_1.

Remark 4.5 If we also assume that each $x \in \mathbf{R}^n$ is contained in sets $R \in \mathcal{B}$ with arbitrarily small diameter then clearly $\mathcal{B}$ is a Busemann-Feller differentiation basis with $\mathcal{B}(x) = \{R : x \in R \in \mathcal{B}\}$.

It is a standard argument that the V_1 property (which is clearly weaker than the CV_1 property) of a B-F basis $\mathcal{B}$ implies that $\mathcal{B}$ differentiates the L_∞ functions, which clearly implies the density property of the basis $\mathcal{B}$. (In fact, as Busemann and Feller proved, differentiating L_∞ is equivalent to the density property). Therefore the minimal density property implies the density property. On the other hand, as we proved the minimal density property of $\mathcal{I}^n$, we have an alternative proof of Saks' strong maximal theorem. (For these definitions and results see e. g. [6] or [7].)

Theorem 4.6

$$\mathrm{MDP} \Leftrightarrow \mathrm{CV}_1.$$

That is, for any class $\mathcal{B}$ of nonempty open bounded subsets of $\mathbf{R}^n$ the minimal density property and the CV_1 property are equivalent.

Proof. According to Remark 4.4, it is enough to prove that $\mathrm{CV}_1 \Rightarrow$ MDP.

Suppose that $\mathcal{R} \subset \mathcal{B}$ covers the measurable $H \subset \mathbf{R}^n$ such that the density of H is d in $\cup\mathcal{R}$ but at most s in any $R \in \mathcal{R}$.

Using the CV_1 property of $\mathcal{B}$ for $\varepsilon = d/2 = \frac{|H|}{2|\cup\mathcal{R}|}$ we get a sequence $R_1, \dots, R_n \in \mathcal{R}$ such that

$$(i) \qquad |\cup\mathcal{R} \setminus \cup_{k=1}^m R_k| \le \varepsilon|\cup\mathcal{R}| = |H|/2$$

and

$$(ii) \qquad \sum_{k=1}^m |R_k| \le C(d/2)|\cup\mathcal{R}|.$$

Since $H \subset \cup\mathcal{R}$, (i) implies that $|H \cap (\cup_{k=1}^m R_k)| \ge |H|/2$. Thus, using that the density of H is at most s in each R_k, we get

$$\frac{|H|}{2} \le |H \cap (\cup_{k=1}^m R_k)| \le \sum_{k=1}^m |H \cap R_k| \le s\sum_{k=1}^m |R_k| \le s\ C(d/2)|\cup\mathcal{R}|.$$

Therefore

$$s \ge \frac{d/2}{C(d/2)},$$

which means that choosing $\rho(d) < \frac{d}{2C(d/2)}$ we get the minimal density property of $\mathcal{B}$. □

Example 4.7 Let $\mathcal{R}$ consist of sets that are the union of an open disc and an open sector with the same centre and twice larger radius.

Then $\mathcal{R}$ is clearly a regular B-F base, so it has several standard nice properties (e.g. weak 1-1 property of the maximal operator, density property, it differentiates L_1 functions).

However $\mathcal{R}$ does not have the minimal density property. Indeed, we can cover an annulus by sets of $\mathcal{R}$ (with the same centre and radius) such that the density of the annulus is arbitrary small in each set.

Therefore

1. The minimal density property is strictly stronger than the density property.
2. The minimal density property and the CV_q properties of a class cannot be proved by using only the standard methods (e.g. properties of the maximal operator).

Remark 4.8 It would be interesting to find a weak sufficient geometrical condition that guarantees the MDP. We could find (see [8]) a quite weak sufficient condition that includes for example the regular convex sets and also the star-shaped sets that contain a ball in their hub with radius at least a fixed constant times the diameter of the set. In fact, we could prove in [8] that this condition implies a Besicovitch type property, which is much stronger than the MDP (or even the CV_∞ property), which shows that the condition is too strong.

Lemma 4.9 *Let $\mathcal{B}$ be a class of nonempty bounded open subsets of $\mathbf{R}^n$ satisfying the minimal density property with the function ρ. Then for any $\varepsilon > 0$ from any sequence $R_1, R_2, \ldots \in \mathcal{B}$ with $|\cup_{i=1}^\infty R_i| < \infty$ one can select a finite subsequence $\tilde{R}_1, \ldots, \tilde{R}_m$ with the following properties:*

$$(i) \qquad |\cup_{k=1}^m \tilde{R}_k| \geq (1-\varepsilon)|\cup_{i=1}^\infty R_i| \qquad \text{and}$$

$$P_1^{\rho(\varepsilon)}: \qquad |\tilde{R}_k \setminus \cup_{j<k}\tilde{R}_j| > \rho(\varepsilon)|\tilde{R}_k| \qquad (k=1,\ldots,m).$$

Proof. We define the subsequence $\tilde{R}_k$ by induction. If $\tilde{R}_1, \ldots, \tilde{R}_{k-1}$ is defined then let $\tilde{R}_k$ be the first element of the sequence (R_i) for which the disjoint part property $\mathrm{P}_1^{\rho(\varepsilon)}$ is satisfied for k. If there is no such R_i then the procedure is finished and $m = k-1$. Thus we get a finite or infinite subsequence $(\tilde{R}_k)$.

The disjoint part property $\mathrm{P}_1^{\rho(\varepsilon)}$ is clearly satisfied for every k, so we have to prove only (i).

Let

$$H = \cup_{i=1}^\infty R_i \setminus \cup_k \tilde{R}_k.$$

Suppose that $x \in H$. Then there exists an index l_x for which $x \in R_{l_x}$. Since R_{l_x} were not chosen in the subsequence $(\tilde{R}_k)$ there exists a k_x for which

$$|R_{l_x} \cap H| \leq |R_{l_x} \setminus \cup_{j<k_x}\tilde{R}_j| \leq \rho(\varepsilon)|R_{l_x}|.$$

Therefore H is covered by $\mathcal{R} = \{R_{l_x} : x \in H\} \subset \mathcal{B}$ such that the density of H is at most $\rho(\varepsilon)$ in each $R \in \mathcal{R}$. Thus, by the minimal density property of $\mathcal{B}$, the density of H in $\cup\mathcal{R}$ is less than ε. Therefore

$$\varepsilon > \frac{|H|}{|\cup_{x\in H} R_{l_x}|} \geq \frac{|H|}{|\cup_{i=1}^\infty R_i|} = \frac{|\cup_{i=1}^\infty R_i \setminus \cup_k \tilde{R}_k|}{|\cup_{i=1}^\infty R_i|}.$$

Hence $|\cup_k \tilde{R}_k| > (1-\varepsilon)|\cup_{i=1}^\infty R_i|$, so, taking m large enough in the case when $(\tilde{R}_k)$ is infinite, we can satisfy (i). □

Notation 4.10 Let $M_{\mathcal{B}}$ denotes the maximal operator corresponding to $\mathcal{B}$, that is

$$M_{\mathcal{B}}(f)(x) = \sup_{x \in R \in \mathcal{B}} \frac{1}{|R|} \int_R |f| \quad \text{if } x \in \cup\mathcal{B}$$

and $M_{\mathcal{B}}(f)(x) = 0$ otherwise.

Theorem 4.11 *Let $1 < p \leq \infty$ and $1/p + 1/q = 1$. If $\mathcal{B}$ has the minimal density property (or the equivalent CV_1 property) and the maximal operator $M_{\mathcal{B}}$ is weak-(p,p) than $\mathcal{B}$ has the CV_q property as well.*

Proof. Let $\mathcal{R} \subset \mathcal{B}$, $|\cup \mathcal{R}| < \infty$ and $\varepsilon > 0$. We can assume that $\mathcal{R} = \{R_1, R_2, \ldots\}$. Then, applying Lemma 4.9, we get a finite subsequence $\tilde{R}_1, \ldots, \tilde{R}_m$ satisfying (i) and $\mathrm{P}_1^{\rho(\varepsilon)}$.

Therefore we only have to prove that the disjoint part property $\mathrm{P}_1^{\rho(\varepsilon)}$ and the weak type (p,p) property of the maximal operator $M_{\mathcal{B}}$ implies that

$$\|\sum_{k=1}^{m} \chi_{\tilde{R}_k}\|_q \leq C(\varepsilon)|\cup \mathcal{R}|^{1/q}.$$

This is essentially proved in [2] in the proof of Proposition 1. (One should only replace 1/2 by $\rho(\varepsilon)$ in that proof). □

Corollary 4.12 *If $\mathcal{B}$ has the MDP (or the equivalent CV_1) then*

$$V_q \Leftrightarrow CV_q \quad (1 \leq q < \infty).$$

Proof. It is proved in [2] that the V_q property of $\mathcal{B}$ and the weak type (p,p) property of the maximal operator $M_{\mathcal{B}}$ are equivalent (if $\frac{1}{p} + \frac{1}{q} = 1$), (in fact, we need only the easy $V_q \Rightarrow$ weak-(p,p) part of this result), hence the non-trivial $V_q \Rightarrow CV_q$ implication follows from Theorem 4.11. □

Corollary 4.13 *The class $\mathcal{I}^n$ has the CV_q property for any $1 \leq q < \infty$; that is, for any $n \in \mathbf{N}$, $1 \leq q < \infty$ and $\varepsilon > 0$ there exists a constant $C(n,q,\varepsilon)$ such that if $\mathcal{R}$ is a family of n-dimensional intervals and $|\cup \mathcal{R}| < \infty$ then there is a finite sequence $R_1, \ldots, R_m \in \mathcal{R}$ such that*

$$\text{(i)} \quad |\cup_{k=1}^m R_k| \geq (1-\varepsilon)|\cup\mathcal{R}| \quad \textit{and} \quad \text{(ii)} \quad \|\sum_{k=1}^{m} \chi_{R_k}\|_q \leq C(n,q,\varepsilon)|\cup\mathcal{R}|^{1/q}. \ \square$$

Remark 4.14 Taking a general Orlicz norm $\|.\|_\Phi$ we can also define the V_Φ and CV_Φ properties by replacing in Definition 4.2 (ii) by $\|\sum_{k=1}^{m}\chi_{R_k}\|_\Phi \leq C\|\chi_{\cup\mathcal{R}}\|_\Phi$. We do not know weather it is always true that if $\mathcal{B}$ has the MDP then $V_\Phi \Leftrightarrow CV_\Phi$.

Acknowledgment. I would like to thank Professors M. Laczkovich and D. Preiss for helpful discussions.

This research was, in part, supported by OTKA grant F 019468. Part of this research was done while the author was at the Mathematical Institute of the Hungarian Academy of Sciences. This note was completed while the author was visiting the University College London having a Royal Society/NATO Postdoctoral Fellowship award.

References

[1] A. Carbery, M. Christ and J. Wright, Multidimensional van der Corput and sublevel set estimates, *preprint.*

[2] A. Córdoba, R. Fefferman, A geometric proof for the strong maximal theorem, *Ann. of Math.* **102** (1975), 95-100.

[3] D. Cruz-Uribe, SFO and C. J. Neugebauer, The structure of the reverse Hölder classes, *Trans. Amer. Math. Soc.* **347** (1995), 2941-2960.

[4] P. Erdős, On extremal problems of graphs and generalized graphs, *Israel J. Math.* **2** (1964), 183-190.

[5] M. de Guzmán, An inequality for the Hardy-Littlewood maximal operator with respect to a product of differentiation bases, *Studia Math.* **49** (1974), 185-194.

[6] M. de Guzmán, *Differentiation of Integrals in* $\mathbf{R}^n$, Springer, Lecture Notes in Mathematics Vol. 481, Berlin, 1975.

[7] M. de Guzmán, *Real Variable Methods in Fourier Analysis*, North-Holland, Mathematics Studies Vol. 46, Amsterdam, 1981.

[8] T. Keleti, A covering property of some classes of sets in $\mathbf{R}^n$, *Acta Univ. Carol. Math. Phys.* **39** (1998), 111-118.

[9] I. Reiman, Über ein problem von K. Zarankiewicz, *Acta Math. Acad. Sci. Hung.* **9** (1958), 269-273.

[10] W. H. Young and G. C. Young, On the existence of a differential coefficient, *Proc. London Math. Soc.*, Ser. 2, **9** (1910), 325-335.

Department of Analysis,
Eötvös Loránd University,
Múzeum krt. 6-8,
H-1088 Budapest,
Hungary
e-mail: elek@cs.elte.hu