Maximal averaging operators: from geometry to boundedness through duality

Alexis Arraz Almirall
Universitat de Barcelona
g2alexisarraz@gmail.com

✉ Christian Cobollo Gómez
Universitat de València
chcogo@alumni.uv.es

Héctor Jardón-Sánchez
University of Oxford
hector.jardonsanchez@keble.ox.ac.uk

Cristina Martín Murillo
Universidad de Extremadura
mmartinue@alumnos.unex.es

✉ Andrés Quilis
Universitat Politècnica de València
anquisan@alumni.uv.es

Pol Ribera Baraut
Universitat de Barcelona
polriberabaraut@gmail.com

Course instructor

Ioannis Parissis
University of the Basque Country
ioannis.parissis@ehu.es

Abstract: We study boundedness of maximal averaging operators through covering properties of their corresponding differentiation bases. This link between analysis and geometry is provided by duality arguments, following the original approach of Córdoba and Fefferman. Ultimately, this provides analogues of the Lebesgue differentiation theorem for certain differentiation bases. We pay special attention to the case of the strong maximal operator, where we prove the strong maximal theorem of Jessen, Marcinkiewicz and Zygmund by means of the previous general approach combined with induction on the dimension and interpolation arguments.

Resumen: Se estudia la acotación de operadores maximales en términos de propiedades de recubrimiento de las correspondientes bases de diferenciación. Esta conexión entre análisis y geometría viene dada por argumentos de dualidad, siguiendo el enfoque original de Córdoba y Fefferman. En último término, lo anterior permite obtener análogos al teorema de diferenciación de Lebesgue para ciertas bases de diferenciación. Se presta especial atención al caso del operador maximal fuerte, para el que se prueba el teorema maximal fuerte de Jessen, Marcinkiewicz y Zygmund combinando el enfoque general anterior con un argumento de inducción en la dimensión y argumentos de interpolación.

Keywords: differentiation basis, duality, strong maximal function, Lebesgue differentiation theorem.

MSC2010: 42B25.

Acknowledgements: The authors would like to thank Ioannis Parissis for his patient and dedicated supervision of the project, as well as the organizers of the IX *Escuela-Taller de Análisis Funcional* held at the University of the Basque Country and the Basque Center for Applied Mathematics in the memory of Bernardo Cascales. The first and sixth authors would like to express their gratitude to María Jesús Carro and Javier Soria for their invitation to the above workshop. The third author was supported by the María Cristina Masaveu Peterson Foundation and the "La Caixa" Foundation, and thanks Pedro Tradacete for the invitation to the workshop.

Reference: Arraz Almirall, Alexis; Cobollo Gómez, Christian; Jardón-Sánchez, Héctor; Martín Murillo, Cristina; Quilis, Andrés; Ribera Baraut, Pol, and Parissis, Ioannis. "Maximal averaging operators: from geometry to boundedness through duality". In: *TEMat monográficos*, 1 (2020): *Artículos de las VIII y IX Escuela-Taller de Análisis Funcional*, pp. 97-111. issn: 2660-6003. url: https://temat.es/monograficos/article/view/vol1-p97.

1. Introduction

The Lebesgue differentiation theorem is a classical result in real analysis (see, for instance, Wheeden and Zygmund's book [10, Chapter 7]) which states that, for every $1 \le p \le \infty$, the collection of open euclidean balls in $\mathbb{R}^n$ differentiate every function in $L^p(\mathbb{R}^n)$ almost everywhere. A more precise statement of this result can be found in theorem 1 below; before its formulation, we introduce some notation and several definitions.

In this article, n will denote a positive natural number and p will be a number on the extended real interval $[1, +\infty]$.

A *differentiation basis* is a family $\mathcal{B}$ consisting of bounded open sets in $\mathbb{R}^n$ whose union is the whole space and which is homothecy invariant, that is, for every $x \in \mathbb{R}^n$, every $\lambda \in \mathbb{R}$, and every $B \in \mathcal{B}$ we have that $B + x, \lambda B \in \mathcal{B}$. It is straightforward to check that the collection of all Euclidean balls in $\mathbb{R}^n$ is a differentiation basis. We will denote this basis by $\mathcal{B}_n$. Two other differentiation bases which we will be using in this text are the cubes in $\mathbb{R}^n$, and the rectangular parallelepipeds (rectangles) in $\mathbb{R}^n$, with sides parallel to the coordinate axes; these will be denoted by $\mathcal{Q}_n$ and $\mathcal{R}_n$, respectively.

Now let $\mathcal{B}$ be a differentiation basis and $\phi : \mathcal{B} \to \mathbb{R}_+$ be a set function. We write

$$\lim_{\substack{B \searrow x \\ B \in \mathcal{B}}} \phi(B)$$

to denote the limit of $\phi(B)$ as the diameter of B tends to 0 and $x \in B$, for sets in the differentiation basis $\mathcal{B}$. When we work with only one differentiation basis we will simply write $\lim_{B \searrow x} \phi(B)$.

With this notation we can state the Lebesgue differentiation theorem in the following way.

Theorem 1 (Lebesgue differentiation theorem). *Let $n \in \mathbb{N}$ and $1 \le p \le \infty$. For every $f \in L^p(\mathbb{R}^n)$ we have that, for almost every $x \in \mathbb{R}^n$,*

$$\lim_{\substack{B \searrow x \\ B \in \mathcal{B}_n}} \frac{1}{|B|} \int_B |f| \to f.$$

One of the purposes of this article is to study conditions under which the Lebesgue differentiation theorem holds when we substitute $\mathcal{B}_n$ by other differentiation bases. We carry out this analysis in detail for the basis $\mathcal{R}_n$ consisting of n-dimensional rectangles.

Arguably, the most important tool in the study of differentiation bases is the corresponding *maximal operator*. This is a sublinear operator that can be associated with every differentiation basis and whose properties are closely related to the respective differentiation properties of the bases. We give the definition below.

Definition 2. Let $\mathcal{B}$ be a differentiation basis. The **maximal operator** associated with $\mathcal{B}$ is defined for all $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ as

$$M_{\mathcal{B}} f(x) := \sup_{\substack{B \in \mathcal{B} \\ x \in B}} \frac{1}{|B|} \int_B f, \qquad x \in \mathbb{R}^n.$$ ◂

It is not hard to see that $M_{\mathcal{B}} f$ is a well-defined measurable function whenever $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$. It is also easy to check that $M_{\mathcal{B}}$ is sublinear on $L^1_{\mathrm{loc}}(\mathbb{R}^n)$:

$$M_{\mathcal{B}}(f+g) \le M_{\mathcal{B}}(f) + M_{\mathcal{B}}(g) \qquad \forall f, g \in L^1_{\mathrm{loc}}(\mathbb{R}^n).$$

An easy consequence of the sublinearity that we will use below is that

$$|M_{\mathcal{B}} f - M_{\mathcal{B}} g| \le M_{\mathcal{B}}(f-g) \qquad \forall f, g \in L^1_{\mathrm{loc}}(\mathbb{R}^n). \tag{1}$$

We briefly discuss the importance of maximal operators in the subject of differentiation bases. For an extensive discussion of the properties of differentiation bases and related differentiation theorems we send the interested readers to De Guzmán's monograph [5].

Let $\mathcal{B}$ be a differentiation basis, let $\varepsilon > 0$, and define the sublinear operator T_ε by means of

$$T_\varepsilon f(x) := \sup_{\substack{B\in\mathcal{B},\ \mathrm{diam}(B)<\varepsilon \\ x\in B}} \frac{1}{|B|}\int_B f, \qquad f \in L^1_{\mathrm{loc}}(\mathbb{R}^n), \quad x \in \mathbb{R}^n.$$

It is clear that theorem 1 is true for $\mathcal{B}$ if $(T_\varepsilon f)_\varepsilon$ converges to f almost everywhere as $\varepsilon \to 0^+$, for every function $f \in L^p(\mathbb{R}^n)$. Let us fix f to be such a function and, for simplicity, let us assume that $1 \le p < \infty$. Then, for every $g \in C_c^\infty(\mathbb{R}^n)$ we use sublinearity in the form of (1) to write

$$|T_\varepsilon f(x) - f(x)| \le |T_\varepsilon(f-g)(x)| + |T_\varepsilon g(x) - g(x)| + |g(x) - f(x)|. \tag{2}$$

Now, in order to prove that $(T_\varepsilon f)_\varepsilon$ converges to f almost everywhere, it is enough to show that

$$\limsup_{\varepsilon\to 0^+} |T_\varepsilon f(x) - f(x)| = 0 \qquad \text{for a.e. } x \in \mathbb{R}^n.$$

In turn, this will follow if we manage to show that for every $\lambda > 0$ we have that

$$|\{x \in \mathbb{R}^n : \limsup_{\varepsilon\to 0^+} |T_\varepsilon f(x) - f(x)| > \lambda\}| = 0. \tag{3}$$

We now fix $\lambda > 0$ and, by estimate (2), we have that

$$\begin{aligned}
&|\{x \in \mathbb{R}^n : \limsup_{\varepsilon\to 0^+} |T_\varepsilon f(x) - f(x)| > \lambda\}| \\
&\quad\le |\{x \in \mathbb{R}^n : \limsup_{\varepsilon\to 0^+} |T_\varepsilon f(x) - T_\varepsilon g(x)| > \lambda/3\}| + |\{x \in \mathbb{R}^n : \limsup_{\varepsilon\to 0^+} |T_\varepsilon g(x) - g(x)| > \lambda/3\}| \\
&\qquad + |\{x \in \mathbb{R}^n : |g(x) - f(x)| > \lambda/3\}| \\
&\quad\le |\{x \in \mathbb{R}^n : M_{\mathcal{B}}(f-g) > \lambda/3\}| + \frac{3^p}{\lambda^p}\|f-g\|^p_{L^p(\mathbb{R}^n)}.
\end{aligned} \tag{4}$$

Passing to the the last line in the estimate above we have used the easily verifiable fact that

$$\limsup_{\varepsilon\to 0^+} |T_\varepsilon g(x) - g(x)| = 0 \qquad \forall g \in C_c^\infty(\mathbb{R}^n)$$

together with the fact that $\sup_{\varepsilon>0} T_\varepsilon(f-g) \le M_{\mathcal{B}}(f-g)$. The condition that we need on $M_{\mathcal{B}}$ in order to deal with the first term in the last line of (4) is given in the following definition.

Definition 3. Let T be a sublinear operator defined on locally integrable functions, and let $1 \le p < \infty$. We say that T is of **weak-type (p, p)** if there exists $C > 0$, depending only upon T, p, and the dimension n, such that

$$|\{x \in \mathbb{R}^n : |Tf(x)| > \lambda\}| \le \frac{C^p}{\lambda^p}\int_{\mathbb{R}^n} |f(x)|^p \,\mathrm{d}x, \qquad \forall f \in L^p(\mathbb{R}^n), \quad \forall \lambda > 0.$$ ◂

From the discussion above about estimate (4), it is easy to show that, if $M_{\mathcal{B}}$ is of weak-type (p, p), then theorem 1 holds for the differentiation basis $\mathcal{B}$ and the index p. Indeed, by (4) and the weak-type property of $M_{\mathcal{B}}$ we have that

$$|\{x \in \mathbb{R}^n : \limsup_{\varepsilon\to 0^+} |T_\varepsilon f(x) - f(x)| > \lambda\}| \le \frac{C^p 3^p}{\lambda^p}\|f-g\|^p_{L^p(\mathbb{R}^n)} + \frac{3^p}{\lambda^p}\|f-g\|^p_{L^p(\mathbb{R}^n)}.$$

Now, for any $\delta > 0$, we can choose $g \in C_c^\infty(\mathbb{R}^n)$ such that $\max(C,1)^p 3^p \lambda^{-p}\|f-g\|^p_{L^p(\mathbb{R}^n)} < \delta/2$, which is possible since $C_c^\infty(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$ for $1 \le p < \infty$, showing that $|\{x \in \mathbb{R}^n : \limsup_{\varepsilon\to 0^+} |T_\varepsilon f(x) - f(x)| > \lambda\}| < \delta$.

The discussion above proves (3), and thus theorem 1, under the weak-type (p, p)-assumption for $M_{\mathcal{B}}$. Therefore, our first goal will be to determine under which circumstances the maximal operator associated to a differentiation basis is of weak-type (p, p). To this end, we establish a geometric characterisation of the weak-type (p, p) property, given in terms of the sets of the differential basis, in section 2. This is used to prove that the basis of cubes $\mathcal{Q}_n$ differentiate every function in $L^1_{\mathrm{loc}}(\mathbb{R}^n)$. In the last section, we study the case of the basis of rectangles $\mathcal{R}_n$, where the associated maximal operator is not of weak type $(1, 1)$. Instead, we prove the so called *strong maximal theorem*, which is the appropriate replacement of the weak-type $(1, 1)$ property for the maximal operator associated with the basis $\mathcal{R}_n$.

In the statement of the theorem below we use the standard notation $\log^+ t := \max(0, \log t)$ for $t > 0$.

Theorem 4 (strong maximal theorem [6]). *The following estimate holds for all* $\lambda > 0$*:*

$$|\{x \in \mathbb{R}^n : M_{\mathcal{R}_n} f(x) > \lambda\}| \leq C_n \int_{\mathbb{R}^n} \frac{|f|}{\lambda}\Big(1 + \Big(\log^+ \frac{|f|}{\lambda}\Big)^{n-1}\Big),$$

where C_n *depends only upon the dimension. It follows that* $\mathcal{R}_n$ *differentiates functions* f *for which*

$$\int_K |f(x)|\big(1 + (\log^+ |f(x)|)^{n-1}\big)\,\mathrm{d}x < \infty \qquad \text{for every compact set } K \subset \mathbb{R}^n.$$

The original proof of this theorem is due to Jessen, Marcinkiewicz, and Zygmund [6]. However, we are not going to reproduce the original analytical proof given there. Instead, we will follow the ideas of Córdoba and Fefferman [2], a geometrical approach using covering properties of a differentiation basis in order to prove boundedness of the corresponding maximal operator. In this context, the precise link between the analytic and geometric statements will be given by duality of suitable function spaces and the adjoint of the (linearised) maximal operator associated with a given differentiation basis.

Note that maximal operators are not linear, but sublinear, so in order to define the adjoint operator we will consider a linear operator T bounded by the maximal operator $M_{\mathcal{B}}$ so that T is a *linearisation* of $M_{\mathcal{B}}$. There are several ways to linearise a maximal operator depending on the differentiation basis $\mathcal{B}$. A useful example of a linearisation technique can be found in the proof of proposition 6.

2. Duality link between analysis and geometry

We start by defining a geometric property for differentiation bases.

Definition 5. Let $1 \leq q \leq \infty$. We say that $\mathcal{B}$ has the **covering property** V_q if there exist $c_1, c_2 > 0$ depending only on $\mathcal{B}, q$ and the dimension such that, for every finite collection $\{B_j\}_{j=1}^N \subset \mathcal{B}$, there exists a finite subcollection $\{\tilde{B}_k\}_{k=1}^M \subseteq \{B_j\}_{j=1}^N$ satisfying

(i) $\displaystyle\Big|\bigcup_{j=1}^N B_j\Big| \leq c_1 \Big|\bigcup_{k=1}^M \tilde{B}_k\Big|,$

(ii) $\displaystyle\Big\|\sum_{k=1}^M \mathbf{1}_{\tilde{B}_k}\Big\|_{L^q(\mathbb{R}^n)} \leq c_2 \Big|\bigcup_{j=1}^N B_j\Big|^{\frac{1}{q}}.$ ◂

Property (i) of the definition above roughly states that we did not lose too much measure when passing to the subcollection, while property (ii) is an L^q-control of the overlap of the sets in the subcollection.

The next proposition establishes the duality between the latter geometric property on the basis $\mathcal{B}$ and the analytical weak-type (p, p) condition of its maximal operator $M_{\mathcal{B}}$.

Proposition 6. *Let* $1 < p < \infty$ *and* $\frac{1}{p} + \frac{1}{p'} = 1$. *The maximal operator* $M_{\mathcal{B}}$ *is of weak-type* (p, p) *if and only if* $\mathcal{B}$ *has the covering property* $V_{p'}$.

Proof. We start by showing necessity. Suppose that $\mathcal{B}$ is a differentiation basis with the covering property $V_{p'}$. Let f be a function in $L^p(\mathbb{R}^n)$ and, for $\lambda > 0$, consider the set $E_\lambda := \{x \in \mathbb{R}^n : M_{\mathcal{B}} f(x) > \lambda\}$. If a point x is in E_λ, by definition we have that there exists a set $B_x \in \mathcal{B}$ containing x such that

$$|B_x| < \frac{1}{\lambda} \int_{B_x} |f(y)|\,\mathrm{d}y. \tag{5}$$

Therefore, we have that the set E_λ is contained in the union of the family $C_\lambda := \{B_x : x \in E_\lambda\} \subset \mathcal{B}$ where the B_x are selected satisfying property (5). It is straightforward to check that this inclusion is actually an equality.

Next, we consider a compact set $K \subset E_\lambda = \bigcup_{x\in E_\lambda} B_x$. By compactness, there exists a finite collection $\{B_j\}_{j=1}^N \subset C_\lambda$ covering K. Now we apply the hypothesis, and there exists a finite subcollection $\{\tilde{B}_k\}_{k=1}^M \subset \{B_j\}_{j=1}^N$ satisfying properties (i) and (ii) from definition 5. By property (i) it follows that

$$|K| \leq c_1 \Big| \bigcup_{k=1}^M \tilde{B}_k \Big|. \tag{6}$$

The sets $\tilde{B}_k$ verify inequality (5) for $k = 1, \dots, M$ because they are chosen from the original collection C_λ. Therefore, we have

$$\Big| \bigcup_{k=1}^M \tilde{B}_k \Big| \leq \sum_{k=1}^M |\tilde{B}_k| \leq \frac{1}{\lambda} \int_{\mathbb{R}^n} \sum_{k=1}^M \mathbf{1}_{\tilde{B}_k}(y)|f(y)|\,\mathrm{d}y. \tag{7}$$

Now, this is the integral of the product of two positive integrable functions, so we can use Hölder's inequality to get

$$\int_{\mathbb{R}^n} \sum_{k=1}^M \mathbf{1}_{\tilde{B}_k}(y)|f(y)|\,\mathrm{d}y \leq \Big\| \sum_{k=1}^M \mathbf{1}_{\tilde{B}_k} \Big\|_{p'} \|f\|_p. \tag{8}$$

Next, combining inequalities (7) and (8) and using property (ii) of the definition of $V_{p'}$, we arrive at

$$\Big| \bigcup_{k=1}^M \tilde{B}_k \Big| \leq \frac{c_2}{\lambda} \Big| \bigcup_{k=1}^M \tilde{B}_k \Big|^{\frac{1}{p'}} \|f\|_p, \tag{9}$$

which, together with (6), implies

$$|K|^{\frac{1}{p}} \leq c_1 \Big| \bigcup_{k=1}^M \tilde{B}_k \Big|^{\frac{1}{p}} \leq \frac{c_1 c_2}{\lambda} \|f\|_p. \tag{10}$$

Finally, since the Lebesgue measure is regular, and we have this inequality for every compact set K contained in E_λ, we conclude that $|E_\lambda|^{\frac{1}{p}} = \sup\{|K| : K \subset E_\lambda,\ K \text{ compact}\}^{\frac{1}{p}} \leq c_1 c_2 \lambda^{-1} \|f\|_p$. Since this holds for every $\lambda > 0$ and for every $f \in L^p(\mathbb{R}^n)$, we get that the maximal operator $M_{\mathcal{B}}$ is of weak-type (p, p).

Now, we are going to see that if $M_{\mathcal{B}}$ is of weak-type (p, p), for $1 < p < \infty$, then $\mathcal{B}$ has the covering property $V_{p'}$. To this end, let us consider a finite collection $\{B_j\}_{j=1}^N \subset \mathcal{B}$. Without loss of generality, we can assume that the sets in this collection are ordered by size in measure, $|B_1| \geq |B_2| \geq \dots \geq |B_N|$; this ordering assumption is just for the sake of specificity. Now, we are going to define a selection algorithm to extract a subcollection $\{\tilde{B}_k\}_{k=1}^M$ satisfying inequalities (i) and (ii) in the definition of $V_{p'}$. Start by taking the biggest set in measure, $\tilde{B}_1 = B_1$. Having chosen $\tilde{B}_1, \dots, \tilde{B}_m$, $m < N$, we choose the next set B to be the largest set in measure from $\{B_{m+1}, \dots, B_N\}$ such that

$$\Big| B \cap \bigcup_{j=1}^m \tilde{B}_j \Big| \leq \frac{1}{2}|B|.$$

This condition tells us that the sets we are selecting do not overlap more than 50 % in measure. Since the original collection was finite, the selection algorithm stops in finitely many steps.

We have selected a subcollection $\{\tilde{B}_k\}_{k=1}^M \subseteq \{B_j\}_{j=1}^N$. Now, we have to use that the sets in this subcollection satisfy certain overlapping properties, and that $M_{\mathcal{B}}$ is weak-type (p, p) by hypothesis, to prove that this subcollection verifies the conditions required in the definition of the covering property $V_{p'}$.

To prove condition (i) of $V_{p'}$, it is enough to prove an inequality of the type

$$\Big| \bigcup_{\substack{B \text{ not} \\ \text{selected}}} B \Big| \leq C_p \Big| \bigcup_{k=1}^M \tilde{B}_k \Big|,$$

with C_p a constant depending only on p. Recall that, if $B \in \{B_j\}_{j=1}^N$ has not been selected, then we have that $|B \cap \bigcup_{k=1}^M \tilde{B}_k| > |B|/2$. Hence, the following inclusions hold,

$$\bigcup_{\substack{B \text{ not} \\ \text{selected}}} B \subseteq \bigcup \Big\{ B : \frac{|B \cap \bigcup_{k=1}^M \tilde{B}_k|}{|B|} > \frac{1}{2} \Big\} \subseteq \Big\{ x \in \mathbb{R}^n : M_{\mathcal{B}}(\mathbf{1}_{\bigcup_{k=1}^M \tilde{B}_k})(x) > \frac{1}{2} \Big\},$$

since, if $x \in B$ with B not selected,

$$M_{\mathcal{B}}(\mathbf{1}_{\bigcup_{k=1}^M \tilde{B}_k})(x) = \sup_{B \ni x, B \in \mathcal{B}} \frac{1}{|B|} \int_B \mathbf{1}_{\bigcup_{k=1}^M \tilde{B}_k}(y)\, \mathrm{d}y = \sup_{B \ni x, B \in \mathcal{B}} \frac{|B \cap \bigcup_{k=1}^M \tilde{B}_k|}{|B|} > \frac{1}{2}.$$

Therefore, using that $M_{\mathcal{B}}$ is of weak-type (p,p), we have that

$$\Big| \bigcup_{\substack{B \text{ not} \\ \text{selected}}} B \Big| \leq \Big| \Big\{ x \in \mathbb{R}^n : M_{\mathcal{B}}(\mathbf{1}_{\bigcup_{k=1}^M \tilde{B}_k})(x) > \frac{1}{2} \Big\} \Big| \leq C_p' 2^p \int_{\mathbb{R}^n} |\mathbf{1}_{\bigcup_{k=1}^M \tilde{B}_k}(y)|^p \, \mathrm{d}y = C_p' 2^p \Big| \bigcup_{k=1}^M \tilde{B}_k \Big|$$

with C_p' a constant depending only on p. Hence,

$$\Big| \bigcup_{j=1}^n B_j \Big| \leq (1 + C_p' 2^p) \Big| \bigcup_{k=1}^M \tilde{B}_k \Big|$$

and we conclude that the subcollection $\{\tilde{B}_k\}_{k=1}^M$ satisfies condition (i) of the covering property $V_{p'}$.

In order to prove condition (ii) of the covering property $V_{p'}$, let us start by defining the collection of sets $\{\tilde{E}_k\}_{k=1}^M$ by setting $\tilde{E}_k := \tilde{B}_k \setminus \bigcup_{j<k} \tilde{B}_j$. It can be easily seen that the sets $\{\tilde{E}_k\}_{k=1}^M$ are pairwise disjoint; furthermore, we have

$$|\tilde{E}_k| \geq \frac{1}{2} |\tilde{B}_k| \qquad \text{and} \qquad \bigcup_{k=1}^M \tilde{E}_k = \bigcup_{k=1}^M \tilde{B}_k.$$

These properties tell us that this new subcollection covers the same space as the one given by the algorithm and that the sets involved have at least half of the measure of the previous ones.

Let us define the following linear and weak-type (p,p) operator

$$T(f)(x) := \sum_{k=1}^M \Big(\frac{1}{|\tilde{B}_k|} \int_{\tilde{B}_k} f(y)\, \mathrm{d}y \Big) \mathbf{1}_{\tilde{E}_k}(x), \qquad x \in \mathbb{R}^n.$$

Observe that for fixed $x \in \mathbb{R}^n$ the sum above collapses to a single term because of the fact that the sets $\{\tilde{E}_k\}_{k=1}^M$ are pairwise disjoint. This readily implies that $T(f)(x) \leq M_{\mathcal{B}}(f)(x)$ for all $f \in L^p(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$. Using Tonelli's theorem we see that

$$T^*(f)(x) = \sum_{k=1}^M \Big(\frac{1}{|\tilde{B}_k|} \int_{\tilde{E}_k} f(y)\, \mathrm{d}y \Big) \mathbf{1}_{\tilde{B}_k}(x), \qquad x \in \mathbb{R}^n,$$

is the adjoint operator of T. Evaluating T^* at $\mathbf{1}_{\bigcup_{k=1}^M \tilde{B}_k}$ and using the properties of the collection $\{\tilde{E}_k\}_{k=1}^M$ above, we get

$$T^*(\mathbf{1}_{\bigcup_{k=1}^M \tilde{B}_k})(x) = \sum_{k=1}^M \Big(\frac{|\tilde{E}_k|}{|\tilde{B}_k|} \Big) \mathbf{1}_{\tilde{B}_k}(x) \geq \frac{1}{2} \sum_{k=1}^M \mathbf{1}_{\tilde{B}_k}(x)$$

for all $x \in \mathbb{R}^n$.

Now we claim that, if a sublinear operator acting on measurable functions in $\mathbb{R}^n$ is of weak-type (p,p) for some $1 < p < \infty$, then for every measurable set $E \subseteq \mathbb{R}^n$ of finite measure we have

$$\int_E |Tf(x)|\, \mathrm{d}x \leq C_{p,n,T} \|f\|_{L^p(\mathbb{R}^n)} |E|^{\frac{1}{p'}}. \tag{11}$$

Assuming this claim for a moment and combining it with the estimate above for T^* we obtain

$$\begin{aligned}\Big|\int_{\mathbb{R}^n}\sum_{k=1}^{M}\mathbf{1}_{\tilde{B}_k}(x)f(x)\,\mathrm{d}x\Big| &\le 2\Big|\int_{\mathbb{R}^n}T^*(\mathbf{1}_{\bigcup_{k=1}^{M}\tilde{B}_k})(x)f(x)\,\mathrm{d}x\Big| = 2\Big|\int_{\bigcup_{k=1}^{M}\tilde{B}_k}T(f)(x)\,\mathrm{d}x\Big| \\ &\le 2C_{p,n}\|f\|_{L^p(\mathbb{R}^n)}\Big|\bigcup_{k=1}^{M}\tilde{B}_k\Big|^{\frac{1}{p'}},\end{aligned}$$

with $C_{p,n}$ a constant depending only on p and the dimension n. Finally, taking the supremum over all functions $f \in L^p(\mathbb{R}^n)$ with $\|f\|_{L^p(\mathbb{R}^n)} \le 1$ and using that the dual space of $L^p(\mathbb{R}^n)$ is $L^{p'}(\mathbb{R}^n)$, we conclude

$$\Big\|\sum_{k=1}^{M}\mathbf{1}_{\tilde{B}_k}\Big\|_{L^{p'}(\mathbb{R}^n)} \le 2C_{p,n}\Big|\bigcup_{k=1}^{M}\tilde{B}_k\Big|^{\frac{1}{p'}},$$

so $\{\tilde{B}_k\}_{k=1}^M$ satisfies condition (ii) of the covering property $V_{p'}$ and $\mathcal{B}$ has the covering property $V_{p'}$, as we wanted to see.

It remains to prove the claim, which is however a straightforward calculation using the layer-cake decomposition [3, Proposition 2.3]. We have for any $\beta > 0$

$$\begin{aligned}\int_E |Tf(x)|\,\mathrm{d}x &= \int_0^\infty |\{x \in E : |Tf(x)| > \lambda\}|\,\mathrm{d}\lambda \le \beta|E| + \int_\beta^\infty \frac{C^p}{\lambda^p}\|f\|^p_{L^p(\mathbb{R}^n)}\,\mathrm{d}\lambda \\ &\le \beta|E| + \frac{C^p}{(p-1)\beta^{p-1}}\|f\|^p_{L^p(\mathbb{R}^n)} \le \frac{C}{(p-1)^{\frac{1}{p}}}\|f\|_{L^p(\mathbb{R}^n)}|E|^{\frac{1}{p'}}\end{aligned}$$

by choosing $\beta^p = \frac{C^p}{p-1}\frac{\|f\|_p^p}{|E|}$. This proves the claim with a constant $C_p \simeq p'$ as $p \to 1^+$. ∎

Before ending this section, let us note a couple of remarks.

Remark 7. The claim in the proof above is only valid for $1 < p < \infty$, as one can see also by inspecting the proof. Furthermore, it can be seen that if $\mathcal{B}$ has the covering property V_∞, then $M_{\mathcal{B}}$ is of weak-type $(1,1)$, but the converse is not true. Another remark that is of some interest is that the claim is actually a characterisation of the weak-type (p,p) for some operator T and $p \in (1,\infty)$. Indeed, assume that (11) is true for p. Then, for $\lambda > 0$, consider the set

$$E_\lambda := \{x \in \mathbb{R}^n : |Tf(x)| > \lambda\}.$$

One needs here some qualitative assumption that will show that $|E_\lambda| < \infty$. This can be made concrete for specific operators T, such as maximal functions, by proving an *a priori* estimate on a nice function f and then extending by density. If one can guarantee that $|E_\lambda| < \infty$, then applying (11) to E yields

$$|E_\lambda|\lambda \le \int_{E_\lambda}|T(f)| \le C_{p,n,T}\|f\|_p|E_\lambda|^{\frac{1}{p'}},$$

which clearly implies the weak-type (p,p) of T when $|E_\lambda| < \infty$.

The principle behind this claim is slightly more general and can be used to define an actual norm on the space $L^{p,\infty}(\mathbb{R}^n)$ for $1 < p < \infty$, which turns these spaces into Banach spaces. For $p = 1$, the space $L^{1,\infty}$ is not normable and only a restricted weaker analogue holds. We refer the interested reader to Grafakos's book [4, Exercise 1.4.14] for further details. ◂

Remark 8. The differentiation basis given by all cubes $\mathcal{Q}_n$ has the covering property V_∞ (this is the well-known Vitali covering lemma [10, Chapter 7]). Hence, we conclude that $\mathcal{Q}_n$ differentiates $L^1(\mathbb{R}^n)$. ◂

3. The strong maximal theorem

In this section we give the proof of theorem 4. First of all, we remember that $\mathcal{R}_n$ is the basis whose elements are open rectangles in $\mathbb{R}^n$ with sides parallel to the coordinate axes.

We begin by describing a negative result.

Proposition 9. *The strong maximal operator $M_{\mathcal{R}_n}$ is not of weak-type* $(1,1)$.

Proof. For simplicity, we provide the details in $\mathbb{R}^2$, but essentially the same construction proves the proposition in any dimension. We are going to see that there is no $c>0$ such that

$$|\{x\in\mathbb{R}^2 : M_{\mathcal{R}_2}f(x)>\lambda\}| \le \frac{c}{\lambda}\int_{\mathbb{R}^2}|f(y)|\,\mathrm{d}y \tag{12}$$

holds for $f\equiv \mathbf{1}_Q$, where $Q=[0,1]^2$. Consider the set $A := \{x\in\mathbb{R}^2 : x_1,x_2>1\}$ and take $x=(x_1,x_2)\in A$; see figure 1. We get that

$$M_{\mathcal{R}_2}f(x) = \sup_{\substack{R\in\mathcal{R}_2\\ R\ni x}}\frac{1}{|R|}\int_R|f(y)|\,\mathrm{d}y = \sup_{\substack{R\in\mathcal{R}_2\\ R\ni x}}\frac{1}{|R|}\int_Q\mathrm{d}y = \sup_{\substack{R\in\mathcal{R}_2\\ R\ni x}}\frac{|Q|}{|R|}\ge\frac{1}{x_1x_2}.$$

Now, for $0<\lambda<1$, let $E_\lambda := \{x\in\mathbb{R}^2 : M_{\mathcal{R}_2}f(x)>\lambda\}$. We have that

$$|E_\lambda| > \left|\left\{x\in\mathbb{R}^2 : x_1x_2<\frac{1}{\lambda}\right\}\right| = \int_{\left\{1<x_1<\frac{1}{\lambda x_2},1<x_2<\frac{1}{\lambda}\right\}}\mathrm{d}x_1\,\mathrm{d}x_2 = \frac{1}{\lambda}\log\frac{1}{\lambda}+1-\frac{1}{\lambda}\simeq\frac{1}{\lambda}\log\frac{1}{\lambda},$$

where in the last step the functions on both sides of $\simeq$ are comparable for $\lambda\in(0,1)$. From here we can conclude that (12) does not hold. Otherwise, we would have that

$$\frac{1}{\lambda}\log\frac{1}{\lambda}\le\frac{c}{\lambda},$$

which is clearly impossible when $\lambda\to 0$.

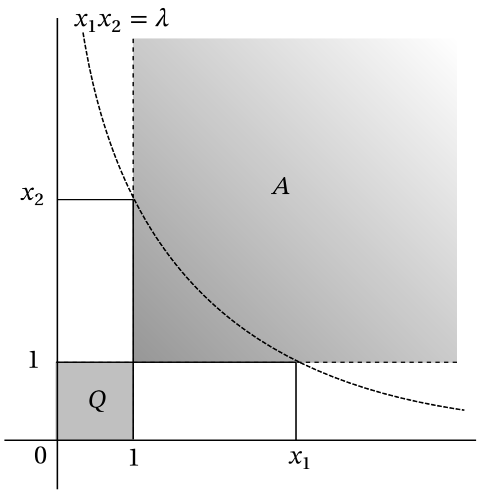

Figure 1: Representation of the function $f\equiv\mathbf{1}_Q$, the set A and the curve x_1x_2.

Thus, the strong maximal operator $M_{\mathcal{R}_2}$ is not of weak-type $(1,1)$. In higher dimensions we just need to work with the higher dimensional unit cube $[0,1]^n$. ∎

In the case of the strong maximal operator, the suitable substitute of the weak $(1,1)$ property is the $L(\log L)^{n-1}$ endpoint estimate of theorem 4. In order to prove this, we will rely on an approach similar to the one outlined in proposition 6, adjusted to the geometry of the basis $\mathcal{R}_n$. The appropriate covering property is given in the following definition.

Definition 10. We say that a differentiation basis $\mathcal{B}$ in $\mathbb{R}^n$ has the **covering property** $V_{\exp,m}$, $m\in\mathbb{N}$, if there exist $c_1,c_2>0$ such that for every finite collection $\{R_j\}_{j=1}^N\subset\mathcal{B}$ there is a finite subcollection $\{\tilde{R}_k\}_{k=1}^M$ such that

(i) $\left|\bigcup_{j=1}^N R_j\right|\le c_1\left|\bigcup_{k=1}^M\tilde{R}_k\right|$,

(ii) there exists $\theta_o(n) > 0$ such that $\displaystyle\int_{\mathbb{R}^n}\left[\exp\Big(\theta\big(\sum_{k=1}^{M}\mathbf{1}_{\tilde{R}_k}\big)^{\frac{1}{m}}\Big)-1\right] \leq \theta c_2\left|\bigcup_{k=1}^{M}\tilde{R}_k\right|$ for every $\theta \in [0,\theta_o(n))$. ◂

The next proposition will be essential in the proof of theorem 4. We give the statement and proof for general bases $\mathcal{B}$ in $\mathbb{R}^n$ although we will only need it for the basis $\mathcal{R}_n$.

Proposition 11. *If the differentiation basis $\mathcal{B}$ has the covering property $V_{\exp,m}$, then there exists $C > 0$ such that*

$$|\{x\in\mathbb{R}^n : M_{\mathcal{B}}f(x)>\lambda\}| \leq C\int_{\mathbb{R}^n}\frac{|f(x)|}{\lambda}\left(1+\log^+\frac{|f(x)|}{\lambda}\right)^m \mathrm{d}x. \tag{13}$$

Proof. We proceed exactly as in the proof where the covering property $V_{p'}$ implied the weak-type (p,p). Consider the set $E_\lambda \coloneqq \{x\in\mathbb{R}^n : M_{\mathcal{B}}f(x)>\lambda\}$. This set can be written as

$$E_\lambda = \bigcup_{x\in E_\lambda} B_x$$

such that, for all $x \in B_x$, we have

$$\frac{1}{|B_x|}\int_{B_x}|f(y)|\,\mathrm{d}y > \lambda.$$

By taking a compact set $K \subset \bigcup_{x\in E_\lambda} B_x$, we can extract a finite subcover such that $K \subset \bigcup_{j=1}^{N} B_j$ with

$$\frac{1}{|B_j|}\int_{B_j}|f(y)|\,\mathrm{d}y > \lambda.$$

By (i) of definition 10, applying the sub-additive property of the measure and using the last inequality,

$$|K| \leq \left|\bigcup_{j=1}^{N}B_j\right| \leq c_1\left|\bigcup_{k=1}^{M}\tilde{B}_k\right| \leq c_1\sum_{k=1}^{M}|\tilde{B}_k| \leq \frac{c_1}{\lambda}\int_{\mathbb{R}^n}\sum_{k=1}^{M}\mathbf{1}_{\tilde{B}_k}(y)|f(y)|\,\mathrm{d}y.$$

Here we need a generalisation of Hölder's inequality matching the exponential norm in property (ii) of definition 10. In order to prove it, note that $\phi(t) = t(1+(\log^+ t)^m)$ is a positive and strictly increasing function in $(0,+\infty)$ and $\phi(0)=0$. Hence, Young's inequality with respect to ϕ guarantees us that

$$st \leq c_{\theta,m}s\left(1+(\log^+ s)^m\right)+\exp(\theta t^{\frac{1}{m}})-1, \tag{14}$$

where $s,t>0$, θ is a small enough positive value and $c_{\theta,m}$ is a constant value that depends on θ and m; a detailed proof of (14) can be found, for example, in the work of Bagby [1]. Setting

$$s \coloneqq \frac{|f(y)|}{\lambda} \qquad\text{and}\qquad t \coloneqq \sum_{k=1}^{M}\mathbf{1}_{\tilde{B}_k}(y),$$

we have, by (14) and (ii) of definition 10 that

$$|K| \leq c_1\left|\bigcup_{k=1}^{M}\tilde{B}_k\right| \leq \frac{c_1}{\lambda}\int_{\mathbb{R}^n}\sum_{k=1}^{M}\mathbf{1}_{\tilde{B}_k}(y)|f(y)|\,\mathrm{d}y \leq c_1c_2\theta\left|\bigcup_{k=1}^{M}\tilde{B}_k\right| + c_1c_{\theta,m}\int_{\mathbb{R}^n}\frac{|f(y)|}{\lambda}\left(1+\left(\log^+\frac{|f(y)|}{\lambda}\right)^m\right)\mathrm{d}y.$$

Notice that

$$c_1\left|\bigcup_{k=1}^{M}\tilde{B}_k\right| \leq c_1c_2\theta\left|\bigcup_{k=1}^{M}\tilde{B}_k\right| + c_1c_{\theta,m}\int_{\mathbb{R}^n}\frac{|f(y)|}{\lambda}\left(1+\left(\log^+\frac{|f(y)|}{\lambda}\right)^m\right)\mathrm{d}y$$

and then

$$\frac{1-c_2\theta}{c_{\theta,m}}\left|\bigcup_{k=1}^{M}\tilde{B}_k\right| \leq \int_{\mathbb{R}^n}\frac{|f(y)|}{\lambda}\left(1+\left(\log^+\frac{|f(y)|}{\lambda}\right)^m\right)\mathrm{d}y.$$

Choosing θ sufficiently small, letting $K \nearrow E_\lambda$ and using the regularity of the Lebesgue measure, we conclude

$$|E_\lambda| \leq C\int_{\mathbb{R}^n}\frac{|f(y)|}{\lambda}\left(1+\left(\log^+\frac{|f(y)|}{\lambda}\right)^m\right)\mathrm{d}y,$$

for $C \equiv C(\theta,m)$. ■

Now, in order to prove our main theorem, it will suffice to show that the differentiation basis $\mathcal{R}_n$ has the covering property $V_{\exp,n-1}$. Note first that we already know the differentiation properties of the strong maximal operator on $\mathbb{R}$, which agrees with the Hardy-Littlewood maximal operator; see remark 8. This sets a first stone on the path for an inductive proof. Indeed, we will prove by induction on the dimension that theorem 4 holds in $\mathbb{R}^n$, with the case $n = 1$, which is the base step of the induction argument, being known to hold true. We will then use the inductive hypothesis, which states that the theorem holds in $\mathbb{R}^{n-1}$, to prove the corresponding covering property $V_{\exp,n-1}$.

With the latter paragraph as a motivation, we introduce two lemmas detailing some precise implications on the boundedness of an operator satisfying (13).

Lemma 12. *Let T be a sublinear operator for which there exists a constant $C > 0$ such that the inequality*

$$|\{x \in \mathbb{R}^n : Tf(x) > \lambda\}| \leq C \int_{\mathbb{R}^n} \frac{|f(x)|}{\lambda} \left(1 + \log^+ \frac{|f(x)|}{\lambda}\right)^m \mathrm{d}x, \tag{15}$$

holds for every measurable function $f : \mathbb{R}^n \to \mathbb{R}$. In addition, assume that $\|T\|_{L^\infty(\mathbb{R}^n)\to L^\infty(\mathbb{R}^n)} \leq 1$. Then, T is weak-type (p, p) for every $p > 1$ and the following inequality holds:

$$|\{x \in \mathbb{R}^n : Tf(x) > \lambda\}| \leq C2^{m+p} \left(\frac{m}{p-1}\right)^m \int_{\mathbb{R}^n} \left(\frac{|f(x)|}{\lambda}\right)^p.$$

Proof. First, fix some $n \in \mathbb{N}$, $p > 1$. For any function f and $\lambda > 0$, define $f_{>\lambda} = f \cdot \mathbf{1}_{>\lambda/2}$, where $\mathbf{1}_{>\lambda/2}$ is the indicator function of the set $\{x \in \mathbb{R}^n : f(x) > \lambda/2\}$. Analogously, define $f_{\leq\lambda/2} = f - f_{>\lambda/2}$. Using the sublinearity of T, it follows that

$$|\{x \in \mathbb{R}^n : Tf(x) > \lambda\}| \leq |\{x \in \mathbb{R}^n : Tf_{>\lambda/2}(x) > \lambda/2\}| + |\{x \in \mathbb{R}^n : Tf_{\leq\lambda/2}(x) > \lambda/2\}|.$$

Note that, since $\|Tf_{\leq\lambda/2}\|_\infty \leq \|T\|_{L^\infty\to L^\infty}\|f_{\leq\lambda/2}\|_\infty \leq \|f_{\leq\lambda/2}\|_\infty \leq \lambda/2$, the latter term in the inequality above equals 0, since an essentially bounded function cannot exceed its essential supremum.

Applying (15) on the surviving term, we get that

$$|\{x \in \mathbb{R}^n : Tf(x) > \lambda\}| \leq 2^m C \int_{\mathbb{R}^n} \frac{|f_{>\lambda/2}(x)|}{\lambda/2} \left(\log \frac{|f_{>\lambda/2}(x)|}{\lambda/2}\right)^m,$$

where we have used the elementary inequality $(1 + t)^m/t^m \leq 2^m$ for $t > 1$. In addition, we can use that, for every $\varepsilon > 0$, the estimate $(\log t)^m / t^{\varepsilon m} \leq 1/\varepsilon^m$ holds to obtain

$$|\{x \in \mathbb{R}^n : Tf(x) > \lambda\}| \leq \frac{C2^m}{\varepsilon^m} \int_{\mathbb{R}^n} \frac{|f_{>\lambda/2}(x)|}{\lambda/2} \left(\frac{|f_{>\lambda/2}(x)|}{\lambda/2}\right)^{\varepsilon m}. \tag{16}$$

Choosing $\varepsilon m = p - 1$ and noting that $f_{\lambda/2} \leq f$ pointwise almost everywhere, we reach at

$$|\{x \in \mathbb{R}^n : Tf(x) > \lambda\}| \leq C2^{m+p} \left(\frac{m}{p-1}\right)^m \int_{\mathbb{R}^n} \left(\frac{|f(x)|}{\lambda}\right)^p,$$

which is the desired weak-type (p, p) estimate. ■

In the following lemma we use the weak-type estimates above to obtain strong-type estimates, with appropriate control over the involved constants.

Lemma 13. *Let T be a sublinear operator for which there exists a constant $C > 0$ for which the inequality*

$$|\{x \in \mathbb{R}^n : Tf(x) > \lambda\}| \leq C_1 \int_{\mathbb{R}^n} \frac{|f(x)|}{\lambda} \left(1 + \log^+ \frac{|f(x)|}{\lambda}\right)^m, \tag{17}$$

holds for every measurable function $f : \mathbb{R}^n \to \mathbb{R}$. In addition, assume that $\|T\|_{L^\infty(\mathbb{R}^n)\to L^\infty(\mathbb{R}^n)} \leq 1$. Then, T is strong-type (p, p) for every $p > 1$ and the following estimate holds:

$$\|T\|_{L^p(\mathbb{R}^n)\to L^p(\mathbb{R}^n)} \leq C(n, m, C_1) \left(\frac{p}{p-1}\right)^{m+1}. \tag{18}$$

Proof. For any fixed $p > 1$, we estimate the L^p-norm of Tf by using the layer-cake decomposition [3, Proposition 2.3], and then considering again the decomposition $f = f_{>\lambda/2} + f_{\leq\lambda/2}$ together with the sublinearity of T, as follows:

$$\|Tf\|^p_{L^p(\mathbb{R}^n)} = p\int_0^\infty \lambda^{p-1}|\{x \in \mathbb{R}^n : Tf(x) > \lambda\}|\,\mathrm{d}\lambda \leq p\int_0^\infty \lambda^{p-1}|\{x \in \mathbb{R}^n : Tf_{>\lambda/2}(x) > \lambda/2\}|\,\mathrm{d}\lambda.$$

Using the estimate (16) from the proof of lemma 12 and switching the order of integration with Fubini's theorem, we get for any $\varepsilon > 0$ that

$$\begin{aligned}\|Tf\|^p_{L^p(\mathbb{R}^n)} &\leq pC_1\frac{2^{m(1+\varepsilon)}}{\varepsilon^m}\int_{\mathbb{R}^n}|f(x)|^{1+\varepsilon m}\Big(\int_0^{2|f(x)|}\lambda^{p-\varepsilon m-2}\,\mathrm{d}\lambda\Big)\,\mathrm{d}x\\ &= pC_1\frac{2^{m(1+\varepsilon)+p}}{\varepsilon^m}\frac{1}{p-\varepsilon m-1}\int_{\mathbb{R}^n}|f(x)|^p\,\mathrm{d}x\end{aligned}$$

as long as $p - 1 > \varepsilon m$. The proof is completed by optimizing in ε under the constraint above, which amounts to the choice $\varepsilon \simeq (p-1)/m$. ∎

We proceed to show the main theorem.

Theorem 14. *Assume that the strong maximal theorem, theorem 4, holds in $\mathbb{R}^{n-1}$. Then, the differentiation basis $\mathcal{R}_n$ has the covering property $V_{\exp,n-1}$.*

Proof. The statement of this theorem essentially proves the inductive step in the proof of theorem 4, in combination with the results presented previously in this section. Let us denote by $\Pi_n(R)$ the projections of $R \in \mathcal{R}_n$ onto the n-th coordinate axis. With this notation we can order the rectangles by choosing one of their sides, say $\Pi_n(R)$. Hence, we have that $|\Pi_n(R_1)| \geq \ldots \geq |\Pi_n(R_N)|$ for $\{R_j\}_{j=1}^N \subset \mathcal{R}_n$. We construct a subcollection $\{\tilde{R}_k\}_{k=1}^M$ as follows.

First, we choose $\tilde{R}_1 := R_1$. Assuming that the rectangles $\{\tilde{R}_1, \ldots, \tilde{R}_k =: R_{k_0}\} \subset \{R_j\}_{j=1}^N$ for some $k_0 < N$ have been selected, we choose $\tilde{R}_{k+1}$ to be the first rectangle $R \in \{R_{k_0+1}, \ldots, R_N\}$ such that either

(19) $$\Big|R \cap \Big(\bigcup_{j\leq k}\tilde{R}_k^*\Big)\Big| \leq \frac{|R|}{2}$$

holds or $R \cap \big(\bigcup_{j\leq k}\tilde{R}_k\big) = \varnothing$. Here, given some $R \in \mathcal{R}_n$, we define R^* to be the rectangle with the same center as R, satisfying $\Pi_n(R^*) = 3\Pi_n(R)$, and having all other sides coinciding with those of R. This selection algorithm terminates in finitely many steps as the original collection was finite.

Let $\{\tilde{R}_j\}_{j=1}^M \subset \{R_j\}_{j=1}^N$, $M \leq N$, denote the subcollection extracted with the previous selection scheme. Our aim is to project our selected rectangles down to $\mathbb{R}^{n-1}$ in such a way that we can exploit the properties of the strong maximal operator in $\mathbb{R}^{n-1}$. For any $R \in \mathcal{R}_n$ and $y \in \mathbb{R}$, let $\Pi_n^{\perp,y}(R)$ denote the slice of the rectangle R by a hyperplane perpendicular to the n-th axis, and crossing the n-th axis at height y; in formulas,

$$\Pi_n^{\perp,y}(R) := \big\{x \in \mathbb{R}^{n-1} : (x,y) \in R\big\}.$$

If a rectangle $R \in \{R_j\}_{j=1}^N$ has not been selected with the previous scheme, then, for some $k \leq N$,

$$\Big|R \cap \Big(\bigcup_{j\leq k}\tilde{R}_j^*\Big)\Big| > \frac{|R|}{2}.$$

Note that R necessarily intersects one of the $\{\tilde{R}_j\}_{j\leq k}$, since otherwise R would have been selected. This implies that $\Pi_n(R) \subseteq \Pi_n(\bigcup_{j\leq k}\tilde{R}_j)$ and so the Π_n projection of the set appearing in the left hand side of the estimate above is $\Pi_n(R)$. Thus, $|\Pi_n(R)|$ can be cancelled from both sides of the estimate, resulting in an analogous sparseness property for the slices

(20) $$\Big|\Pi_n^{\perp,y}(R) \cap \Big(\bigcup_{j\leq k}\Pi_n^{\perp,y}(\tilde{R}_j^*)\Big)\Big| > \frac{|\Pi_n^{\perp,y}(R)|}{2}.$$

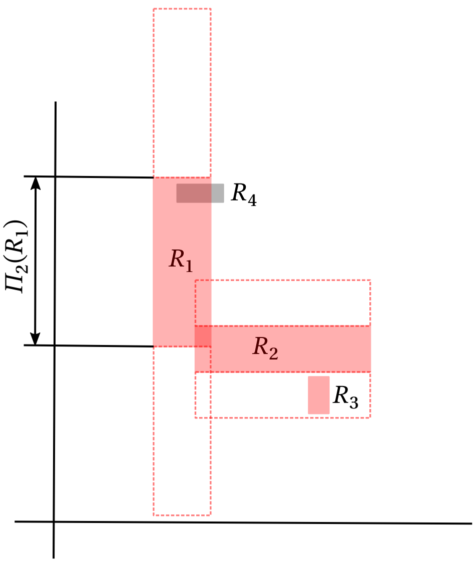

Figure 2: An example of application of the presented selection scheme in $\mathbb{R}^2$. Rectangles are ordered by the size of their vertical projection. Those rectangles shaded in red are the selected ones. The dotted lines show some of the tripled extensions $R \mapsto R^*$.

Estimate (20) must be understood as the statement that the $(n-1)$-dimensional average of the function $\mathbf{1}_{\bigcup_{j\le k}\Pi_n^{\perp,y}(\tilde{R}_j^*)}$ on the $(n-1)$*-dimensional rectangle* $\Pi_n^{\perp,y}(R)$ is big. Thus, remembering the definition of the strong maximal operator as a supremum, we get that, for every $x \in \Pi_n^{\perp,y}(R)$,

$$M_{\mathcal{R}_{n-1}}\big[\mathbf{1}_{\bigcup_{j=1}^{M}\Pi_n^{\perp,y}(\tilde{R}_j^*)}\big](x) \ge M_{\mathcal{R}_{n-1}}\big[\mathbf{1}_{\bigcup_{j\le k}\Pi_n^{\perp,y}(\tilde{R}_j^*)}\big](x) \ge \frac{1}{|\Pi_n^{\perp,y}(R)|}\Big|\Pi_n^{\perp,y}(R)\cap\Big(\bigcup_{j\le k}\Pi_1^{y}(\tilde{R}_j^*)\Big)\Big| > \frac{1}{2}.$$

The estimate above clearly holds also for any $x \in \bigcup_{j=1}^{M}\tilde{R}_j$. Thus, we have proved

$$\textbf{(21)}\qquad \Pi_n^{\perp,y}\Big(\bigcup_{j=1}^{N}R_j\Big) = \bigcup_{j=1}^{N}\Pi_n^{\perp,y}(R_j) \subseteq \Big\{x \in \mathbb{R}^{n-1} \,:\, M_{\mathcal{R}_{n-1}}\big[\mathbf{1}_{\bigcup_{j=1}^{M}\Pi_n^{\perp,y}(\tilde{R}_j^*)}\big](x) > \frac{1}{2}\Big\}.$$

Now we can proceed on showing that our selection scheme actually extracts a subcollection of rectangles satisfying the $V_{\exp,n-1}$ property, namely conditions (i) and (i) of definition 10. Since by assumption $M_{\mathcal{R}_{n-1}}$ satisfies the strong maximal theorem in $\mathbb{R}^{n-1}$, lemma 12 implies that $M_{\mathcal{R}_{n-1}}$ is (say) weak-type $(2,2)$. This and estimate (21) imply that

$$\Big|\Pi_n^{\perp,y}\Big(\bigcup_{j=1}^{N}R_j\Big)\Big| \le C_1\Big|\Pi_n^{\perp,y}\Big(\bigcup_{j=1}^{M}\tilde{R}_j\Big)\Big|$$

for some constant $C_1 > 0$ depending only upon the dimension. Integrating for $y \in \mathbb{R} = \Pi_n(\mathbb{R}^n)$ proves property (i) of $V_{\exp,n-1}$.

In order to prove condition (ii), we may take the following exponential expansion for some $\theta > 0$,

$$\textbf{(22)}\qquad \int_{\mathbb{R}^{n-1}}\Big[\exp\Big(\theta\Big(\sum_{j=1}^{M}\mathbf{1}_{\tilde{R}_j}\Big)^{\frac{1}{n-1}}\Big)-1\Big] = \int_{\mathbb{R}^{n-1}}\sum_{\tau=1}^{\infty}\frac{\theta^\tau}{\tau!}\Big(\sum_{j=1}^{M}\mathbf{1}_{\tilde{R}_j}\Big)^{\frac{\tau}{n-1}} = \sum_{\tau=1}^{\infty}\frac{\theta^\tau}{\tau!}\Big\|\sum_{j=1}^{M}\mathbf{1}_{\tilde{R}_j}\Big\|_{L^{\frac{\tau}{n-1}}(\mathbb{R}^{n-1})}^{\frac{\tau}{n-1}},$$

by an application of the monotone convergence theorem to the partial sums of the exponential series. Again, we may get a control for the overlap in the right hand side of the estimate above from the $(n-1)$-dimensional properties of the strong maximal operator. For any $k \le M$, it follows after the selection rule of equation (19) that

$$\Big|\Pi_n^{\perp,y}(\tilde{R}_k)\cap\Big(\bigcup_{j<k}\Pi_n^{\perp,y}(\tilde{R}_j)\Big)\Big| \le \Big|\Pi_n^{\perp,y}(\tilde{R}_k)\cap\Big(\bigcup_{j<k}\Pi_n^{\perp,y}(\tilde{R}_j^*)\Big)\Big| \le \frac{1}{2}|\Pi_n^{\perp,y}(\tilde{R}_k)|,$$

for every $y \in \mathbb{R}$. In order to simplify the notation, fix any $y \in \mathbb{R}$ and write $I_k := \Pi_n^{y,\perp}(\tilde{R}_k)$, for $k \le M$. In this way, the previous inequality turns into the $(n-1)$-dimensional *sparseness* property

$$\Big|I_k \cap \Big(\bigcup_{j<k} I_j\Big)\Big| \le \frac{1}{2}|I_k|, \qquad k \le M. \tag{23}$$

Consider the disjoint increment sets $E_k := I_k \setminus \bigcup_{j<k} I_j$ for $k \le M$ and define the auxiliary linear operator

$$Tf := \sum_{j=1}^{M} \left(\frac{1}{|I_j|} \int_{I_j} f \right) \mathbf{1}_{E_j} \le M_{\mathcal{R}_{n-1}} f.$$

It follows that T has the same boundedness properties as $M_{\mathcal{R}_{n-1}}$ and so, combining the assumption with lemma 13 gives the estimate

$$\|T\|_{L^q(\mathbb{R}^{n-1}) \to L^q(\mathbb{R}^{n-1})} \le C_n (q')^{n-1}, \qquad \frac{1}{q} + \frac{1}{q'} = 1, \quad q \ge 2. \tag{24}$$

Following the path of the proof of proposition 6, we get that its adjoint is given by the formula

$$T^* f = \sum_{j=1}^{M} \left(\frac{1}{|I_j|} \int_{E_j} f \right) \mathbf{1}_{I_j}, \qquad T^*(\mathbf{1}_{\bigcup_{k=1}^M I_k}) = \sum_{j=1}^{M} \frac{|E_j|}{|I_j|} \mathbf{1}_{I_j}.$$

Note that (23) implies that $|E_k| \ge \frac{1}{2}|I_k|$, so that $T^*(\mathbf{1}_{\bigcup_{k=1}^M I_k}) > \frac{1}{2}\sum_{j=1}^M \mathbf{1}_{I_j}$. This fact, together with (24), allows us to estimate, for every integer $1 \le p < \infty$,

$$\begin{aligned} \left(\int_{\mathbb{R}^{n-1}} \Big| \sum_{k=1}^{M} \mathbf{1}_{I_k} \Big|^p \right)^{\frac{1}{p}} &\le 2 \left(\int_{\mathbb{R}} \Big| T^* \Big(\mathbf{1}_{\bigcup_{k=1}^M I_k} \Big) \Big|^p \right)^{\frac{1}{p}} \le \|T^*\|_{L^{p'}(\mathbb{R}^{n-1}) \to L^{p'}(\mathbb{R}^{n-1})} \Big\| \mathbf{1}_{\bigcup_{k=1}^M I_k} \Big\|_{L^p(\mathbb{R}^{n-1})} \\ &\le C_n p^{n-1} \Big| \bigcup_{k=1}^{M} I_k \Big|^{\frac{1}{p}}, \end{aligned} \tag{25}$$

for some constant $C_n > 0$ depending only upon the dimension. Note that the estimate for $p = 1$ is a straightforward application of (23), without appealing to (24).

We can now complete the proof of condition (ii) of $V_{\exp,n-1}$. Remember that from (22) we have

$$\int_{\mathbb{R}^2} \left[\exp\left(\theta \Big(\sum_{k=1}^{M} \mathbf{1}_{\tilde{R}_k} \Big)^{\frac{1}{n-1}} - 1 \right) \right] \le \left(\sum_{\tau=1}^{2(n-1)} + \sum_{\tau=2n-1}^{\infty} \right) \frac{\theta^\tau}{\tau!} \int_{\mathbb{R}^{n-1}} \Big| \sum_{k=1}^{M} \mathbf{1}_{I_k} \Big|^{\frac{\tau}{n-1}} =: I + II.$$

Now for I, since $\tau/(n-1) \le 2$, we can use Hölder's inequality together with (25) for $p = 2$ to get

$$I \le C_n' \theta e \Big| \bigcup_{k=1}^{M} I_k \Big|,$$

with C_n' only depending on the dimension. For II we will use the asymptotic estimates for $\tau!$ provided by Stirling's formula [8, Exercise 12-22] in the form

$$\lim_{j \to \infty} \frac{j! e^j}{j^j \sqrt{2\pi j}} = 1.$$

This, together with (25) for $p = \tau/(n-1)$, yields the estimate

$$II \le C_n \sum_{\tau=2n-1}^{\infty} \frac{(\theta e)^\tau}{\tau^\tau \sqrt{\tau}} \left(\frac{\tau}{n-1} \right)^\tau \Big| \bigcup_{k=1}^{M} I_k \Big| \le C_n'' \theta \Big| \bigcup_{k=1}^{M} I_k \Big|$$

provided that $\theta < \theta_o(n)$ is sufficiently small; the constant $C_n'' > 0$ above depends only upon the dimension. Summing the estimates for I and II completes the proof of property (ii) of $V_{\exp,n-1}$, and thus the proof of the theorem. ∎

Proof of Theorem 4. We can now put together the full proof of theorem 4, which is by way of induction on the dimension n. For $n = 1$ the theorem holds because of remark 8, namely because $\mathcal{R}_1$ is the basis of intervals of $\mathbb{R}$ and $M_{\mathcal{R}_1}$ is just the Hardy-Littlewood maximal operator. So assume that the theorem holds for $M_{\mathcal{R}_{n-1}}$ in $\mathbb{R}^{n-1}$. Then, theorem 14 tells us that the basis of n-dimensional rectangles has the property $V_{\exp,n-1}$ on $\mathbb{R}^n$, and proposition 11, applied for $m = n - 1$, yields the conclusion of theorem 4 in $\mathbb{R}^n$. The inductive step and thus the proof of the main estimate of the theorem is complete. In order to show that $\mathcal{R}_n$ differentiates functions which are locally in the space $L(\log L)^{n-1}$, namely the second conclusion of the theorem, one follows the argument on p. 3 of section 1, replacing the weak (p, p) type of $M_{\mathcal{B}}$ with the main estimate just proved for $M_{\mathcal{R}_n}$. We omit the details. ∎

4. Concluding remarks

In lemma 13 we presented a particular case of a more general interpolation theorem called the *Marcinkiewicz interpolation theorem*; see Stein's book [9, § I.4], for example. Whenever we have certain boundedness properties of an operator at two particular endpoints $1 \le p < q \le \infty$, we can use interpolation arguments to recover boundedness for every other r in between p and q. In our particular case, we had strong-type (∞, ∞) properties plus the strong maximal theorem, which implies we can use any $p > 1$ as a weak-type (p, p) endpoint. For a more detailed exposition of interpolation theorems for operators acting on different Banach spaces and applications to several problems in harmonic analysis see for example Grafakos's book [4].

This text is intended to be an introduction to the study of maximal operators given by differentiation bases beyond the one consisting of Euclidean balls or cubes in $\mathbb{R}^n$. Several other bases are of interest and give rise to intriguing problems in harmonic analysis. For example one can consider the bases of rectangles in $\mathbb{R}^2$ with short side of length $\delta \ll 1$ and long side of length 1, whose longest side points lie in a given finite set of directions $V \subset \mathbb{S}^1$. If these directions are uniformly distributed on $\mathbb{S}^1$, then this basis gives rise to the so-called *Kakeya maximal function*, an object which is central in one of the main conjectures in modern harmonic analysis. The study of the basis of rectangles with sides parallel to the coordinate sides is a toy model, allowing the development of geometric and combinatorial arguments which are suitable for these more general bases. As mentioned in the introduction, the interested reader could consult De Guzmán's work [5] for a thorough discussion on the theory of differentiation bases and the analysis of several different approaches for their study, together with corresponding conjectures.

References

[1] Bagby, Richard J. "Maximal functions and rearrangements: some new proofs". In: *Indiana University Mathematics Journal* 32.6 (1983), pp. 879–891. issn: 0022-2518. https://doi.org/10.1512/iumj.1983.32.32060.

[2] Córdoba, Antonio J. and Fefferman, Robert A. "A geometric proof of the strong maximal theorem". In: *Annals of Mathematics. Second Series* 102.1 (1975), pp. 95–100. issn: 0003-486X. https://doi.org/10.2307/1970976.

[3] Duoandikoetxea, Javier. *Fourier analysis.* Vol. 29. Graduate Studies in Mathematics. Translated and revised from the 1995 Spanish original by David Cruz-Uribe. Providence: American Mathematical Society, 2001. isbn: 978-0-8218-2172-5.

[4] Grafakos, Loukas. *Classical Fourier analysis.* 3rd ed. Graduate Texts in Mathematics 249. New York: Springer, 2014. https://doi.org/10.1007/978-1-4939-1194-3.

[5] Guzmán, Miguel de. "Differentiation of integrals in $\mathbf{R}^n$". In: *Measure theory.* Conference on Measure Theory (Oberwolfach, 1975). Lecture Notes in Mathematics 541. 1976, pp. 181–185. https://doi.org/10.1007/BFb0081051.

[6] Jessen, Børge; Marcinkiewicz, Józef, and Zygmund, Antoni. "Note on the differentiability of multiple integrals". In: *Fundamenta Mathematicae* 25 (1935), pp. 217–234. issn: 0016-2736. https://doi.org/10.4064/fm-25-1-217-234.

[7] Moriyón, Roberto. *The degenerate complex Monge-Ampère equation.* PhD thesis. Princeton University, 1979.

[8] Schröder, Bernd S. W. *Mathematical analysis. A concise introduction.* Hoboken: John Wiley & Sons, 2008. isbn: 978-0-470-10796-6.

[9] Stein, Elias M. *Singular integrals and differentiability properties of functions.* Princeton Mathematical Series 30. Princeton: Princeton University Press, 1971. isbn: 978-0-691-08079-6.

[10] Wheeden, Richard L. and Zygmund, Antoni. *Measure and integral. An introduction to real analysis.* Pure and Applied Mathematics 43. New York: Marcel Dekker, 1977. isbn: 978-0-8247-6499-9.