Illinois Journal of Mathematics
Volume 56, Number 3, Fall 2012, Pages 759–793
S 0019-2082

MOVING AVERAGES IN THE PLANE

LAURENT MOONENS AND JOSEPH M. ROSENBLATT

Abstract. We study the almost everywhere behavior of the maximal operator associated to moving averages in the plane, both for Lebesgue derivatives and ergodic averages. We show that the almost everywhere behavior of the maximal operator associated to a sequence of moving rectangles $v_i + Q_i$, with $(0,0) \in Q_i$, depends both on the way the rectangles are moved by v_i and the structure of the rectangles (Q_i) as a partially ordered set.

Given a sequence of rectangles $Q_0, Q_1, \dots$ in the plane whose lower left corner is the origin, a result by Stokolos [15, Lemma 1] guarantees that the associated maximal operator M defined by

$$Mf(x) = \sup_i \frac{1}{|Q_i|} \mathbb{1}_{Q_i} * |f|(x)$$

does not satisfy an inequality of weak type $(1,1)$ in case the sequence Q_0^*, $Q_1^*, \dots$ of dyadic approximations of $Q_0, Q_1, \dots$, contains arbitrarily many incomparable rectangles. Using general principles instead of Stokolos' direct method, we show that [15, Lemma 1] implies that M cannot satisfy an inequality of weak type in any Orlicz space $\phi(L)$ with $\phi(x) = o(x \log^+(x))$, showing immediately that the a.e. finiteness of Mf does not occur in those spaces. This is the object of Section 1.1.

On the other hand, rewriting conveniently Stokolos' proof (which is our Lemma 4) allows us to obtain similar results in case the sequence $Q_0, Q_1, \dots$ is shifted by a sequence of vectors $v_1, v_2, \dots$. In particular, we demonstrate

Received June 30, 2011; received in final form August 25, 2011.
Laurent Moonens is a *Chargé de Recherches* of the *Fonds de la Recherche scientifique—FNRS* (Belgium) and has been partially supported by the Belgian American Educational Foundation (B.A.E.F.), New Haven, CT. Joseph Rosenblatt recognizes the partial support of NSF through DMS 0555905.
2010 *Mathematics Subject Classification.* Primary 42B25. Secondary 28D05.

how the almost everywhere behavior of the maximal operator M' defined by

$$M'f(x) = \sup_i \frac{1}{|Q_i|} \mathbb{1}_{v_i+Q_i} * |f|(x),$$

certainly relies on the the sequence (v_i), but even more so on the nature of (Q_i). We show this in Section 1.2.

According to a transfer lemma (Lemma 42, see Appendix B), part of those results can be formulated in the ergodic context (see Sections 2.1 and 2.2). A result by Bellow, Jones and Rosenblatt [2] then allows us to study the behavior of the latter maximal operator on $L_p(X)$ with $1 \leq p \leq \infty$.

In the paper, we use material about Orlicz spaces as well as transfer results; those elements will be found in the Appendices.

1. Moving averages in the plane: The differentiation context

We study first the behaviour of standard averaging in $\mathbb{R}^2$, both in the differentiation and ergodic contexts.

1.1. Standard averages for differentiation. In the sequel, we will call *standard rectangle* any rectangle of the form $[0,\alpha] \times [0,\beta]$ in $\mathbb{R}^2$.

Definition 1. Two standard rectangles Q_1 and Q_2 in $\mathbb{R}^2$ are called *incomparable*, and we write $Q_1 \nsim Q_2$, in case neither $Q_1 \subseteq Q_2$ nor $Q_2 \subseteq Q_1$.

Moreover, a family $\mathbf{Q}$ of standard rectangles is called

- *independent* in case $Q_1 \nsim Q_2$ holds for any distinct $Q_1, Q_2 \in \mathbf{Q}$;
- *dependent* in case there exists distinct elements $Q_1, Q_2 \in \mathbf{Q}$ where $Q_1 \nsim Q_2$ does not hold;
- *a chain* in case $\mathbf{Q}$ is totally ordered by inclusion.

We also introduce the following definition. In the sequel, we let $\mathbb{N} = \{0, 1, 2, \dots\}$ denote the set of all natural numbers, and we let $\mathbb{N}^* = \mathbb{N} \setminus \{0\}$.

Definition 2. A sequence $\mathbf{Q}$ of standard rectangles is said to have *infinite width* in case for every $k \in \mathbb{N}$, there exists integers $1 \leq i_1 < i_2 < \cdots < i_k$ for which $\{Q_{i_j} : 1 \leq j \leq k\}$ is independent. It is said to have *finite width* in case it is not of infinite width.

According to Dilworth's theorem (see Dilworth [5, Theorem 1.1, p. 161]), we have the following alternative.

Lemma 3. *A sequence* $\mathbf{Q}$ *of standard rectangles has either infinite width, and it is not a finite union of chains, or it has finite width and it is a finite union of chains.*

Proof. Clearly, if $\mathbf{Q}$ has infinite width then it cannot be a finite union of chains. If $\mathbf{Q}$ does not have infinite width, there exists $k \in \mathbb{N}^*$ such that for any family of indices $1 \leq i_1 < i_2 < \cdots < i_k$ we have $Q_{i_j} \sim Q_{i_l}$ for some $1 \leq j < l \leq k$; that is to say that every subset of $\mathbf{Q}$ counting k elements is

dependent. According to Dilworth's theorem (see Dilworth [5, Theorem 1.1, p. 161]), $\mathbf{Q}$ can be written as an union of finitely many chains. □

The key for the study of this case is a result by Stokolos [15, Lemma 1]. We make some simplifications in Stokolos' proof and use some notation that we believe helps motivate and explain the result. In particular, we will be using the Rademacher functions (r_i) which are defined on $[0,1)$ as follows. Let $r_1(x) = \mathbb{1}_{[0,1/2)}(x)$. Then inductively let $r_i(x) = r_{i-1}(2x \mod 1)$ for $i \geq 2$. The Rademacher functions (r_i) form an IID sequence on $[0,1]$ taking on only the values 0 and 1 (see Zygmund [16, p. 6]).

LEMMA 4 (Stokolos). *We assume that we have an* independent *family of standard dyadic rectangles* $\{Q_i : 1 \leq i \leq k\}$. *There are Lebesgue measurable sets* Θ *and* Y *in* $[0,1]^2$ *with the following properties*:

(i) $\Theta \subset Y$;
(ii) $|Y| \geq \frac{1}{4} k 2^k |\Theta|$;
(iii) *for all* $x \in Y$, *there exists* $u_0 = u_0(x) \in [0,1]^2$ *and* $\nu_0 = \nu_0(x)$, $1 \leq \nu_0 \leq k$ *such that* $x \in u_0 + Q_{\nu_0}$ *and*

$$\frac{|(u_0 + Q_{\nu_0}) \cap \Theta|}{|Q_{\nu_0}|} \geq \frac{1}{2^{k-1}}.$$

Proof. We may assume that $Q_i = [0, 2^{-m_i}] \times [0, 2^{-n_i}]$ where $1 \leq n_1 < \cdots < n_k$ and $m_1 > \cdots > m_k \geq 1$. Let Θ be such that the characteristic function 1_Θ is given by

$$\mathbb{1}_\Theta(\xi, \eta) = \prod_{i=1}^{k} r_{m_i}(\xi) \prod_{j=1}^{k} r_{n_j}(\eta).$$

For $1 \leq \nu \leq k$, we also define Y_ν similarly by

$$\mathbb{1}_{Y_\nu}(\xi, \eta) = \prod_{i=\nu}^{k} r_{m_i}(\xi) \prod_{j=1}^{\nu} r_{n_j}(\eta).$$

We let $Y = \bigcup_{\nu=1}^{k} Y_\nu$.

Now we can compute $|\Theta|$ as follows:

$$|\Theta| = \int\int 1_\Theta(\xi, \eta)\, d\xi\, d\eta = \int\int \prod_{i=1}^{k} r_{m_i}(\xi) \prod_{j=1}^{k} r_{n_j}(\eta)\, d\xi\, d\eta = \frac{1}{2^{2k}},$$

using the independence of the Rademacher functions. It is worthwhile to also observe that Θ consists of $2^{-2k} 2^{m_1} 2^{n_k}$ dyadic rectangles of side length 2^{-m_1} parallel to the ξ-axis and side length 2^{-n_k} parallel to the η-axis.

Similarly, we compute:

$$|Y_\nu| = \int\int 1_{Y_\nu}(\xi, \eta)\, d\xi\, d\eta = \int\int \prod_{i=\nu}^{k} r_{m_i}(\xi) \prod_{j=1}^{\nu} r_{n_j}(\eta)\, d\xi\, d\eta = \frac{1}{2^{k+1}}$$

using the independence of the Rademacher functions. In this case, Y_ν consists of $2^{-(k+1)}2^{m_\nu}2^{n_\nu}$ rectangles of side length 2^{-m_ν} parallel to the ξ-axis and side length 2^{-n_ν} parallel to the η-axis. It follows that $\Theta \subset Y_\nu$ and $\frac{|\Theta|}{|Y_\nu|} = \frac{2}{2^k}$ for all $\nu = 1, \dots, k$.

In order to estimate $|Y|$, we first estimate, for each $2 \le \nu \le k$, the measure of the set E_ν defined by

$$E_\nu = Y_\nu \cap \bigcup_{p=1}^{\nu-1} Y_p \subseteq Y_\nu.$$

To this purpose, we first notice that given $x = (\xi, \eta) \in E_\nu$ we have, by definition of Y_ν and Y_p, $1 \le p \le \nu - 1$:

$$\prod_{i=\nu}^{k} r_{m_i}(\xi) \prod_{j=1}^{\nu} r_{n_j}(\eta) = 1 \quad \text{and}$$

$$\prod_{i=p_0}^{k} r_{m_i}(\xi) \prod_{j=1}^{p_0} r_{n_j}(\eta) = 1 \quad \text{for some } 1 \le p_0 \le \nu - 1.$$

As $p_0 \le \nu - 1$, this yields in particular

$$\prod_{i=\nu-1}^{k} r_{m_i}(\xi) \prod_{j=1}^{\nu} r_{n_j}(\eta) = 1.$$

We hence get

$$|E_\nu| \le \int\int \prod_{i=\nu-1}^{k} r_{m_i}(\xi) \prod_{j=1}^{\nu} r_{n_j}(\eta)\, d\xi\, d\eta = \frac{1}{2^{k+2}} = \frac{1}{2}|Y_\nu|,$$

so that one finally estimates

$$|Y| = |Y_1| + \sum_{\nu=2}^{k} |Y_\nu \setminus E_\nu| = |Y_1| + \sum_{\nu=2}^{k} \big(|Y_\nu| - |E_\nu|\big) \ge \frac{1}{2} \sum_{\nu=1}^{k} |Y_\nu| = \frac{1}{4} k 2^k |\Theta|.$$

This inequality gives us the basic estimates on the sizes of the sets Θ and Y that we wanted. We now need to verify that the proportionality facts hold for suitable rectangles. To that purpose, fix $x \in Y$ and denote by $\nu_0 = \nu_0(x)$ the smallest integer $1 \le \nu \le k$ for which $x \in Y_\nu$ holds. If we take the $2^{-m_{\nu_0}} \times 2^{-n_{\nu_0}}$ rectangle $R_0 = R_0(x) \subset Y_{\nu_0}$ which contains x, then we have $R_0 = u_0 + Q_{\nu_0}$ for some $u_0 = u_0(x) \in [0,1]^2$. Because $\Theta \subset Y_\nu$, we have

$$\frac{|R_0 \cap \Theta|}{|R_0|} = \frac{|Y_\nu \cap \Theta|}{|Y_\nu|} = \frac{|\Theta|}{|Y_\nu|} = \frac{1}{2^{k-1}}.$$

Given what we have observed already, the only part of this that needs some explanation is the first equality. But this equality holds because Y_ν is a union

of $N_\nu = 2^{-(k+1)}2^{m_\nu}2^{n_\nu}$ disjoint $2^{-m_\nu}\times 2^{-n_\nu}$ rectangles whose intersections with Θ have the same measure, i.e. $|R_0\cap\Theta|$. So

$$\frac{|Y_\nu\cap\Theta|}{|Y_\nu|}=\frac{N_\nu|R_0\cap\Theta|}{N_\nu|R_0|}=\frac{|R_0\cap\Theta|}{|R_0|}.$$

The proof is complete. □

REMARK 5. Let us see what this construction becomes in the simplest case, where the m_i and n_j are changing by 1 each time the index i or j changes by 1. That is, $m_i = k-i+1$ and $n_i = i$ for $i = 1,\dots,k$, and

$$R_i=\left[0,2^{-m_i}\right]\times\left[0,2^{-n_i}\right]=\left[0,2^{-(k-i+1)}\right]\times\left[0,2^{-i}\right].$$

Then Θ consists of $2^{-2k}2^{m_1}2^{n_k}=2^{-2k}2^k2^k=1$ rectangle, namely $[0,2^{-k}]\times[0,2^{-k}]$. Also, each Y_ν consists of $2^{-(k+1)}2^{m_\nu}2^{n_\nu}=2^{-(k+1)}2^{k-\nu+1}2^\nu=1$ rectangle, namely $[0,2^{-\nu}]\times[0,2^{-(k-\nu+1)}]$. In particular, $Y_\nu = R_{k-\nu+1}$.

REMARK 6. With slightly more work, the above construction can be carried out so that Y is in a preassigned rectangle $[0,\delta]\times[0,\delta]$ and so has as small diameter as we would like.

In the sequel, we fix a sequence $\mathbf{Q}=(Q_i)$ of standard rectangles. We associate to $\mathbf{Q}$ a differentiation basis

$$\mathscr{B}_{\mathbf{Q}}=\left\{x+Q_i : x\in\mathbb{R}^2, i\in\mathbb{N}^*\right\},$$

and we let, for $x\in\mathbb{R}^2$,

$$\mathscr{B}_{\mathbf{Q}}(x)=\{R\in\mathscr{B}_{\mathbf{Q}} : R\ni x\}.$$

We define the maximal differentiation operator $D_{\mathbf{Q}}$ on $L^1(\mathbb{R}^2)$ by

$$D_{\mathbf{Q}}f(x)=\sup\left\{\frac{1}{|R|}\int_R f : R\in\mathscr{B}_{\mathbf{Q}}(x)\right\}.$$

We introduce the following definition.

DEFINITION 7. Given a sequence $\mathbf{Q}=(Q_i)$ of standard rectangles, we denote by $\mathbf{Q}^*=(Q_i^*)$ the sequence of standard dyadic rectangles obtained in the following way: for each $i\in\mathbb{N}$, $Q_i^*=[0,2^{-m_i}]\times[0,2^{-n_i}]$ $(m_i,n_i\in\mathbb{Z})$ is the standard dyadic rectangle of minimal measure containing Q_i.

As Stokolos observes in [15, Remark 1, p. 106], Lemma 4 provides a lower estimate on the measure of the level sets of the maximal differentiation operator when working with a sequence $\mathbf{Q}$ of standard rectangle such that $\mathbf{Q}^*$ has infinite width.

Using Lemma 4, we can now prove the following result. Let $\Phi_0(t):=t(1+\log_+ t)$ (see Appendix A and Example 41).

COROLLARY 8. *Let $\mathbf{Q}$ be a family of standard dyadic rectangles. If $\mathbf{Q}$ has infinite width, then for each $\lambda\geq 1$ there exists a sequence of functions $(f_k)\subseteq L^\infty_+(\mathbb{R}^2)$ satisfying the following conditions: for each $k\in\mathbb{N}^*$,*

(i) *f_k vanishes outside* $[0,1]^2$;
(ii) $\int_{\mathbb{R}^2} \Phi_0(|f_k|)\,dx \geq \frac{1}{2}k\|f_k\|_1$;
(iii) $|\{x \in \mathbb{R}^2 : D_{\mathbf{Q}} f_k \geq \lambda\}| \geq \frac{1}{2\lambda}\int_{\mathbb{R}^2} \Phi_0(|f_k|)\,dx$ *for k sufficiently large*;
(iv) *for any Orlicz function Φ satisfying $\Phi = o(\Phi_0)$ at ∞ we have*

$$\lim_{k\to\infty} \frac{\int_{\mathbb{R}^2} \Phi_0(|f_k|)\,dx}{\int_{\mathbb{R}^2} \Phi(C|f_k|)\,dx} = \infty,$$

for each $C > 0$.

Proof. Take $k \in \mathbb{N}^*$. Since $\mathbf{Q}$ has infinite width, find indices $1 \leq i_1 < i_2 < \cdots < i_k$ for which $\{Q_{i_j} : 1 \leq j \leq k\}$ is independent. For each $1 \leq j \leq k$, let $\tilde{Q}_j = Q_{i_j}$ and choose sets Θ_k and Y_k in $[0,1]^2$ associated to $\tilde{Q}_j$, $1 \leq i \leq k$ according to Lemma 4. Let $f_k = 2^{k-1}\lambda \mathbb{1}_{\Theta_k}$. It is clear that f_k is supported in $[0,1]^2$, and (i) is proved.

To prove (ii), begin by observing that

$$\begin{aligned}\int_{\mathbb{R}^2} \Phi_0\big(|f_k|\big)\,dx &= |\Theta_k| 2^{k-1}\lambda\big[1 + \log\lambda + (k-1)\log 2\big] \\ &= 2^{-k-1}\lambda\big[1 + \log\lambda + (k-1)\log 2\big].\end{aligned}$$

On the other hand, we have

$$\|f_k\|_1 = 2^{k-1}\lambda|\Theta_k| = 2^{-k-1}\lambda.$$

We hence get:

$$\int_{\mathbb{R}^2} \Phi_0\big(|f_k|\big)\,dx = \big[1 + \log\lambda + (k-1)\log 2\big]\|f_k\|_1 \geq \frac{1}{2}k\|f_k\|_1.$$

To show (iii), observe now that given $x \in Y_k$, we have

$$D_{\mathbf{Q}} f_k(x) \geq \frac{1}{|\tilde{Q}_{\nu_0}|}\int_{u_0+\tilde{Q}_{\nu_0}} 2^{k-1}\lambda \mathbb{1}_{\Theta_k}(y)\,dy = 2^{k-1}\lambda\frac{|(u_0+\tilde{Q}_{\nu_0})\cap\Theta_k|}{|\tilde{Q}_{\nu_0}|} \geq \lambda,$$

where u_0 and ν_0 are associated to x by Lemma 4. In particular, we get

$$\big|\big\{x \in \mathbb{R}^2 : D_{\mathbf{Q}} f(x) \geq \lambda\big\}\big| \geq |Y_k| \geq \frac{1}{4}k2^k|\Theta_k| = \frac{1}{2}k2^{-k-1}.$$

This finally yields, for k sufficiently large:

$$\begin{aligned}\big|\big\{x \in \mathbb{R}^2 : D_{\mathbf{Q}} f_k(x) \geq \lambda\big\}\big| &\geq \frac{1}{2\lambda}\lambda k 2^{-k-1} \\ &= \frac{1}{2\lambda}\frac{k}{1 + \log\lambda + (k-1)\log 2}\int_{\mathbb{R}^2} \Phi_0\big(|f_k|\big)\,dx \\ &\geq \frac{1}{2\lambda}\int_{\mathbb{R}^2} \Phi_0\big(|f_k|\big)\,dx,\end{aligned}$$

and the proof of (iii) is complete.

To prove (iv), fix an Orlicz function Φ satisfying $\lim_{t\to\infty}\frac{\Phi(t)}{\Phi_0(t)}=0$ and let $C>1$ be a constant. Observing that for any $k\in\mathbb{N}^*$ we have

$$\frac{\int_{\mathbb{R}^2}\Phi(C|f_k|)\,dx}{\int_{\mathbb{R}^2}\Phi_0(|f_k|)\,dx}=\frac{\Phi(2^{k-1}\lambda C)}{\Phi_0(2^{k-1}\lambda)}=\frac{\Phi(2^{k-1}\lambda C)}{\Phi_0(2^{k-1}\lambda C)}\cdot\frac{\Phi_0(2^{k-1}\lambda C)}{\Phi_0(2^{k-1}\lambda)}.$$

On the other hand, using the doubling condition Φ_0 satisfies (see Appendix A and Example 41) it is easily shown that $\Phi_0(2^{k-1}\lambda C)/\Phi_0(2^{k-1}\lambda)$ is bounded by a constant depending only on λ as k goes to ∞; it hence follows that

$$\lim_{k\to\infty}\frac{\int_{\mathbb{R}^2}\Phi(C|f_k|)\,dx}{\int_{\mathbb{R}^2}\Phi_0(|f_k|)\,dx}=0,$$

and the proof is complete. □

Assume that $\mathbf{Q}$ is a sequence of standard rectangles such that $\mathbf{Q}^*$ has infinite width. It now follows immediately from Corollary 8 that the maximal operator $D_{\mathbf{Q}}$ cannot be of weak type (Φ,Φ) in case Φ is an Orlicz function satisfying $\Phi=o(\Phi_0)$ at ∞, where Φ_0 is the Orlicz function in Example 41. Let us first illustrate this by showing that the maximal operator $D_{\mathbf{Q}}$ cannot satisfy a weak $(1,1)$ inequality. To this purpose, we proceed towards a contradiction, and assume there exists a constant $C>0$ independent of k such that for any $k\in\mathbb{N}^*$ we have

$$\big|\big\{x\in\mathbb{R}^2:D_{\mathbf{Q}}f_k(x)>1/2\big\}\big|\le 2C\|f_k\|_1. \tag{1}$$

We let $\tilde{\mathbf{Q}}=(\tilde{Q}_i)$ be defined by $\tilde{Q}_i=\frac{1}{2}Q_i$. Observing that $\tilde{\mathbf{Q}}^*$ has infinite width (as one easily checks it), we apply Corollary 8 to $\tilde{\mathbf{Q}}^*$ and $\lambda=3$, and denote by (f_k) the associated sequence of functions. Using the inequality

$$D_{\mathbf{Q}}f_k(x)\ge\frac{1}{4}D_{\tilde{\mathbf{Q}}^*}f_k(x),$$

valid for each $x\in\mathbb{R}^2$ and each $k\in\mathbb{N}^*$, we get

$$\big|\big\{x\in\mathbb{R}^2:D_{\tilde{\mathbf{Q}}^*}f_k(x)>2\big\}\big|\le\big|\big\{x\in\mathbb{R}^2:D_{\mathbf{Q}}f_k(x)>1/2\big\}\big|\le 2C\|f_k\|_1.$$

On the other hand, according to Corollary 8(ii), (iii), we have for k sufficiently large:

$$\begin{aligned}\big|\big\{x\in\mathbb{R}^2:D_{\tilde{\mathbf{Q}}^*}f_k(x)>2\big\}\big|&\ge\big|\big\{x\in\mathbb{R}^2:D_{\tilde{\mathbf{Q}}^*}f_k(x)\ge 3\big\}\big|\\&\ge\frac{1}{6}\int_{\mathbb{R}^2}\Phi_0\big(|f_k|\big)\,dx\ge\frac{1}{12}k\|f_k\|_1,\end{aligned}$$

which is contradictory to the previous estimate.

Going back to the general case, let Φ be an Orlicz function satisfying $\Phi=o(\Phi_0)$ at ∞. We now show that $D_{\mathbf{Q}}$ cannot satisfy an inequality of weak type

(Φ, Φ). To see it, we proceed again towards a contradiction: assume that there would exist a constant $C > 0$ such that

$$\left|\left\{x \in \mathbb{R}^2 : D_{\mathbf{Q}} f(x) > \lambda\right\}\right| \leq \int_{\mathbb{R}^2} \Phi\left(\frac{C|f|}{\lambda}\right) dx, \tag{2}$$

holds for every $f \in L^{\Phi}_{+}(\mathbb{R}^2)$ and every $\lambda > 0$. Let us keep the notations used before and let (f_k) be the sequence of functions associated to $\tilde{\mathbf{Q}}^*$ and $\lambda = 3$ by Corollary 8. We would then have, according to Corollary 8(iii), for k sufficiently large:

$$\begin{aligned}\frac{1}{6}\int_{\mathbb{R}^2} \Phi_0\big(|f_k|\big)\, dx &\leq \left|\left\{x \in \mathbb{R}^2 : D_{\tilde{\mathbf{Q}}^*} f_k(x) \geq 3\right\}\right| \\ &\leq \left|\left\{x \in \mathbb{R}^2 : D_{\mathbf{Q}} f_k(x) > \frac{1}{2}\right\}\right| \leq \int_{\mathbb{R}^2} \Phi\big(2C|f_k|\big)\, dx.\end{aligned}$$

This would contradict, for k sufficiently large, Corollary 8(iv); hence $D_{\mathbf{Q}}$ cannot satisfy an inequality of weak type as above.

Given a sequence of standard rectangles $\mathbf{Q}$, we are actually interested in studying the behavior of the maximal operator $M^*_{\mathbf{Q}}$ associated to $\mathbf{Q}$—which, unlike $D_{\mathbf{Q}}$, remains relevant in the study of moving averages—defined by

$$M_{\mathbf{Q}} f(x) = \sup_{i \in \mathbb{N}^*} \frac{1}{|Q_i|} \mathbb{1}_{Q_i} * f(x),$$

for a.e. $x \in \mathbb{R}^2$.

Remark 9. In fact, the maximal operators $D_{\mathbf{Q}}$ and $M_{\mathbf{Q}}$ are *distributionally equivalent*; more precisely, for each $f \in L^1_+(\mathbb{R}^2)$ and each $\lambda > 0$, we have:

(i) $|\{x \in \mathbb{R}^2 : M_{\mathbf{Q}} f(x) > \lambda\}| \leq |\{x \in \mathbb{R}^2 : D_{\mathbf{Q}} f(x) > \lambda\}|$;
(ii) $|\{x \in \mathbb{R}^2 : D_{\mathbf{Q}} f(x) > \lambda\}| \leq \sum_{p,q=0}^{1} |\{x \in \mathbb{R}^2 : M_{\mathbf{Q}}(f \circ T_{p,q})(x) > \frac{\lambda}{4}\}|$;

where for $0 \leq p, q \leq 1$, the operator $T_{p,q} : \mathbb{R}^2 \to \mathbb{R}^2$ is defined by $T_{p,q}(\xi, \eta) = ((-1)^p \xi, (-1)^q \eta)$.

As (i) follows immediately from the inequality $D_{\mathbf{Q}} f(x) \geq M_{\mathbf{Q}} f(x)$, valid for each $f \in L^1_+(\mathbb{R}^2)$ and each $x \in \mathbb{R}^2$, only (ii) has to be explained. To that purpose, fix $f \in L^1_+(\mathbb{R}^2)$ and observe that for $x \in \mathbb{R}^2$ with $D_{\mathbf{Q}} f(x) > \lambda$, there exists $i \in \mathbb{N}^*$ and $u \in \mathbb{R}^2$ such that

$$x \in u + Q_i \quad \text{and} \quad \frac{1}{|Q_i|}\int_{u+Q_i} f > \lambda. \tag{3}$$

It follows from (3) that we have

$$\sum_{p,q=0}^{1} \frac{1}{|Q_i|}\int_{x+T_{p,q}Q_i} f = \frac{1}{|Q_i|}\int_{\bigcup_{p,q=0}^{1} x+T_{p,q}Q_i} f \geq \frac{1}{|Q_i|}\int_{u+Q_i} f > \lambda;$$

this yields $|Q_i|^{-1}\int_{x+T_{p,q}Q_i} f > \lambda/4$ for some $0 \le p, q \le 1$; that is

$$\frac{1}{|Q_i|}\int_{T_{|p-1|,|q-1|}x-Q_i} f \circ T_{|p-1|,|q-1|} > \frac{\lambda}{4}.$$

We conclude from this discussion that

$$\big\{x \in \mathbb{R}^2 : D_{\mathbf{Q}} f(x) > \lambda\big\} \subseteq \bigcup_{p,q=0}^{1} T_{p,q}\bigg\{x \in \mathbb{R}^2 : M_{\mathbf{Q}}(f \circ T_{p,q})(x) > \frac{\lambda}{4}\bigg\}.$$

The maximal operator $M_{\mathbf{Q}}$ has the following a.e. behaviour.

THEOREM 10. *Let* $\mathbf{Q}$ *be a sequence of standard cubes. We have the following properties*:

(i) *if* $\mathbf{Q}^*$ *has finite width and if* $f \in L^1_+(\mathbb{R}^2)$, *then* $M_{\mathbf{Q}} f < \infty$ *almost everywhere on* $\mathbb{R}^2$;

(ii) *if* $\mathbf{Q}^*$ *has infinite width and if* Φ *is an Orlicz function satisfying* $\Phi = o(\Phi_0)$ *at* ∞, *then there exists* $f \in L^\Phi_+(\mathbb{R}^2)$ *for which* $M_{\mathbf{Q}} f = \infty$ *almost everywhere on* $\mathbb{R}^2$; *in particular, there exists* $f \in L^1_+(\mathbb{R}^2)$ *for which* $M_{\mathbf{Q}} f = \infty$ *almost everywhere on* $\mathbb{R}^2$;

(iii) *in general, for any* $\mathbf{Q}$, *if* $f \in L\log L_+(\mathbb{R}^2)$, *then* $M_{\mathbf{Q}} f < \infty$ *almost everywhere on* $\mathbb{R}^2$; *in particular, if* $f \in L^p(\mathbb{R}^2)$ *for some* $1 < p \le \infty$, *then we have* $M_{\mathbf{Q}} f < \infty$, *a.e. in* $\mathbb{R}^2$.

In order to prove Theorem 10, we need to make a few observations.

REMARK 11. If Φ is an Orlicz function, then one easily shows from the Lebesgue dominated convergence theorem that if $(\rho_k) \subseteq C_c^\infty(\mathbb{R}^n)$ is a regularizing sequence (see Appendix B for a precise definition), then

$$\|f - \rho_k * f\|_\Phi \to 0, \quad k \to \infty.$$

In particular, $C_c^\infty(\mathbb{R}^n)$ is dense in $L^\Phi(\mathbb{R}^n)$.

REMARK 12. Let Φ be an Orlicz function. Before proving Theorem 10, it should be readily noticed that for any measurable set $A \subseteq \mathbb{R}^2$ having finite Lebesgue measure, Jensen's inequality applied to the normalized Lebesgue measure on A yields the following inequality for each $f \in L^\Phi(\mathbb{R}^2)$:

$$\Phi\big(|A|^{-1}|\mathbb{1}_A * f|\big) \le \Phi\bigg(\frac{1}{|A|}\int_A \big|f(\cdot - \xi)\big|\,d\xi\bigg) \le \frac{1}{|A|}\int_A \Phi\big(\big|f(\cdot - \xi)\big|\big)\,d\xi.$$

It hence follows from an application of Fubini's theorem that the operator

$$L^\Phi\big(\mathbb{R}^2\big) \to L^\Phi\big(\mathbb{R}^2\big), \qquad f \mapsto |A|^{-1}\mathbb{1}_A * f,$$

has strong type (Φ, Φ).

Proof of Theorem 10. To prove (i), assume that $\mathbf{Q}^*$ has finite width; in particular, it is a finite union of chains $\mathbf{Q}_1^*, \dots, \mathbf{Q}_l^*$. It follows from Zygmund [16] (yet in the ergodic case) that for each $1 \le j \le l$, $M_{\mathbf{Q}_j^*}$ is of type $(1,1)$. Using the inequality

$$M_{\mathbf{Q}} f(x) \le 4 M_{\mathbf{Q}^*} f(x),$$

valid for each measurable f and each $x \in \mathbb{R}^2$, we then observe that for each $\lambda > 0$ and each $f \in L^1_+(\mathbb{R}^2)$ we have

$$\{x \in \mathbb{R}^2 : M_{\mathbf{Q}} f(x) > 4\lambda\} \subseteq \{x \in \mathbb{R}^2 : M_{\mathbf{Q}^*} f(x) > \lambda\}.$$

As one easily checks, also, that

$$\{x \in \mathbb{R}^2 : M_{\mathbf{Q}^*} f(x) > \lambda\} \subseteq \bigcup_{j=1}^{l} \{x \in \mathbb{R}^2 : M_{\mathbf{Q}_j^*} f(x) > \lambda\},$$

one finishes the proof of (i) by applying [13, Theorem 3].

If (ii) did not hold, we would deduce from Stein [14, Theorem 3] that $M_{\mathbf{Q}}$, and hence $D_{\mathbf{Q}}$, is of weak type (Φ, Φ), contradicting Remark 2.

Finally, it follows from Jessen, Marcinkiewicz and Zygmund [8, Theorems B or 4] that $D_{\mathbf{Q}}$ satisfies an inequality of weak type in either $L \log L(\mathbb{R}^2)$ or $L^p(\mathbb{R}^2)$ in case $p < \infty$, or an inequality of strong type in $L^\infty(\mathbb{R}^2)$ in case $p = \infty$. We then infer from Remark 9 that $M_{\mathbf{Q}}$ satisfies a similar inequality. If $f \in L^\infty_+(\mathbb{R}^2)$, then the same conclusion follows from Bellow and Jones [3, Corollary 1]. □

Remark 13. Observe that (ii) is the analogue of Stokolos [15, Theorem 1, part 2] for the maximal operator $M_{\mathbf{Q}}$. We obtained it here using general principles instead of a direct computation.

Remark 14. Let $\mathbf{Q}$ be a family of standard rectangles, define $\tilde{\mathbf{Q}} = (\tilde{Q}_i)$ by $\tilde{Q}_i = \frac{1}{2} Q_i$, and assume that the following property is satisfied:

(∗) for each $k \in \mathbb{N}^*$, there exists a subfamily $\tilde{\mathbf{Q}}_k^*$ of $\tilde{\mathbf{Q}}^*$ containing exactly k distinct rectangles with *equal areas*.

Then, a much simpler argument (due to B. Reznick) shows that $M_{\mathbf{Q}}$ cannot satisfy a weak $(1,1)$ inequality. One sees it using the following:

Claim 15. *For each even $k \in \mathbb{N}^*$, we have $|\bigcup \tilde{\mathbf{Q}}_k^*| \ge \frac{1}{3} k \alpha_k$, where α_k denotes the value of the area of each rectangle in $\tilde{\mathbf{Q}}_k^*$.*

Proof. Fix an even $k \in \mathbb{N}^*$, write $k = 2l$ and begin by choosing an ordering $\tilde{\mathbf{Q}}_k^* = (\tilde{Q}_k^1, \dots, \tilde{Q}_k^k)$ making their sides $c_1, \dots, c_k$ along the x-axis into a decreasing sequence. In particular, we have $c_{i+1} \le c_i/2$ for each $1 \le i \le k-1$, and, in general, for $1 \le i \le i+j \le k$, $c_{i+j} \le 2^{-j} c_i$; in particular, we have

$$|\tilde{Q}_k^i \cap \tilde{Q}_k^{i+j}| \le 2^{-j} \alpha_k.$$

It is obvious that

$$\left|\bigcup_{i=1}^{k}\tilde{Q}_k^i\right| \geq \left|\bigcup_{r=1}^{l}\tilde{Q}_k^{2r}\right|.$$

On the other hand, a simple induction argument then shows that

$$\left|\bigcup_{r=1}^{l}\tilde{Q}_k^{2r}\right| \geq \bigcup_{r=1}^{l}\left|\tilde{Q}_k^{2r}\right| - \sum_{1\leq s<t\leq l}\left|\tilde{Q}_k^{2s}\cap\tilde{Q}_k^{2t}\right| = l\alpha_k - \sum_{t=2}^{l}\sum_{s=1}^{t-1}2^{-2(t-s)}\alpha_k.$$

As one easily computes

$$\sum_{t=2}^{l}\sum_{s=1}^{t-1}2^{-2(t-s)} = \frac{1}{3}\sum_{t=2}^{l}\left(1-4^{-t}\right) \leq \frac{l}{3},$$

one finally gets

$$\left|\bigcup_{r=1}^{l}\tilde{Q}_k^{2r}\right| \geq l\alpha_k - \frac{1}{3}l\alpha_k = \frac{1}{3}(2l)\alpha_k,$$

which yields the result. □

Keeping the notations of the preceding proof, fix an integer $k\geq 2$, let $A_k = \bigcap\tilde{\mathbf{Q}}_k^*$ and observe that $|A_k|\leq 2^{-k}\alpha_k$. Define $f_k\in L^1_+(\mathbb{R}^2)$ by $f_k = |A_k|^{-1}\mathbb{1}_{A_k}$. Given $1\leq i\leq k$, observe that for any $x\in\tilde{Q}_k^i\setminus A_k$, we have

$$\frac{1}{|\tilde{Q}_k^i|}\mathbb{1}_{\tilde{Q}_k^i} * f_k(x) = \frac{1}{|A_k|}\frac{1}{|\tilde{Q}_k^i|}\int_{\tilde{Q}_k^i}\mathbb{1}_{A_k}(x-\xi)\,d\xi = \frac{1}{|A_k|}\frac{|\tilde{Q}_k^i\cap(x-A_k)|}{|\tilde{Q}_k^i|} = \frac{1}{\alpha_k},$$

for in such a case, we have $|\tilde{Q}_k^i\cap(x-A_k)| = |A_k|$. Consequently, we get

$$\begin{aligned}
&\left|\left\{x\in\mathbb{R}^2 : M_{\tilde{\mathbf{Q}}^*}f_k(x)\geq\frac{1}{\alpha_k}\right\}\right| \\
&\quad\geq \left|\bigcup_{i=1}^{k}(\tilde{Q}_k^i\setminus A_k)\right| \geq \left|\bigcup_{i=1}^{k}\tilde{Q}_k^i\right| - |A_k| \\
&\quad\geq \left(\frac{k}{3}-\frac{1}{2^k}\right)\alpha_k \geq \frac{1}{3}(k-1)\alpha_k = \frac{1}{3}(k-1)\frac{\|f_k\|_1}{\frac{1}{\alpha_k}}.
\end{aligned}$$

It then follows from the inequality $M_{\mathbf{Q}}f_k(x)\geq\frac{1}{4}M_{\tilde{\mathbf{Q}}^*}f_k(x)$, valid for each $x\in\mathbb{R}^2$, that we have

$$\left|\left\{x\in\mathbb{R}^2 : M_{\mathbf{Q}}f_k(x)>\frac{3}{\alpha_k}\right\}\right| \geq (k-1)\frac{\|f_k\|_1}{(\frac{3}{\alpha_k})};$$

hence $M_{\mathbf{Q}}$ cannot be of weak type $(1,1)$.

REMARK 16. Given a sequence $\mathbf{Q}$ of standard rectangles satisfying $\operatorname{diam}(Q_i)\to 0$ as $i\to\infty$, we can make the following observations:

(i) for any $\mathbf{Q}$, the sequence of functions defined by
$$x \mapsto \frac{1}{|Q_i|}\chi_{Q_i} * f(x),$$
converges in L^1-norm to f;
(ii) if $\mathbf{Q}^*$ has finite width, then the sequence of functions defined in (i) converges pointwise a.e. to f for all $f \in L^1(\mathbb{R}^2)$;
(iii) if $\mathbf{Q}^*$ has infinite width, then the sequence of functions defined in (i) fails to converge pointwise a.e. for some $f \in L^1(\mathbb{R}^2)$ and hence for a generic $f \in L^1(\mathbb{R}^2)$;
(iv) in general, for any $\mathbf{Q}$, the sequence of functions defined in (i) converges pointwise a.e. to f for all $f \in L\log L(\mathbb{R}^2)$.

1.2. Moving averages for differentiation. In order to deal with moving averages, we shall need to prove a result about the independence of translates of the Rademacher functions.

Given $t_0 \in [0,1)$ and a function f defined on $[0,1)$, we let $\tau_{t_0} f$ be defined on $[0,1)$ by $\tau_{t_0} f(t) = f((t+t_0) \mod 1)$.

LEMMA 17. *For any sequence $(t_i) \subseteq [0,1)$, the translated Rademacher functions $\tau_{t_i} r_i$, $i \in \mathbb{N}^*$ form an IID sequence on $[0,1)$.*

Proof. It is clear that the functions $\tau_{t_i} r_i$, $i \in \mathbb{N}^*$ are identically distributed for each sequence $(t_i) \subseteq [0,1)$ because each one is a characteristic function on a set of Lebesgue measure $1/2$.

To prove that they form an independent sequence, we begin by proving the following identity.

CLAIM 18. *For any sequence $(t_i) \subseteq [0,1)$, we have*
$$\big|\{\tau_{t_1} r_1 = 1, \dots, \tau_{t_k} r_k = 1\}\big| = \frac{1}{2^k}. \tag{4}$$

Proof. We proceed by induction on k. We first notice that the result is trivial for $k = 1$. Assuming it has been proved for $k \le l-1$ with $l \ge 2$, we observe that by invariance under translation, it suffices to prove the identity (4) for a sequence (t_i) satisfying $t_1 = 0$. But then what we have to show is that
$$\int_0^{1/2} \mathbb{1}_{\{\tau_{t_2} r_2 = 1\}}(t) \cdots \mathbb{1}_{\{\tau_{t_l} r_l = 1\}}(t)\, dt = \int_0^{1/2} \tau_{t_2} r_2(t) \cdots \tau_{t_l} r_l(t)\, dt = \frac{1}{2^l}.$$
For $1 \le i \le l-1$, let $s_i = 2t_{i+1}$, and observe that we have, for $0 \le t < 1/2$:
$$\tau_{t_i} r_i(t) = r_i\big[(t+t_i) \mod 1\big] = r_{i-1}\big[(2t+2t_i) \mod 1\big] = \tau_{s_{i-1}} r_{i-1}(2t).$$
Using the substitution $s = 2t$, we hence get
$$\int_0^{1/2} \tau_{t_2} r_2(t) \cdots \tau_{t_l} r_l(t)\, dt = \frac{1}{2}\int_0^1 \tau_{s_1} r_1(s) \cdots \tau_{s_{l-1}} r_{l-1}(s)\, ds = \frac{1}{2} \cdot \frac{1}{2^{l-1}} = \frac{1}{2^l},$$

using the induction hypothesis, which proves the claim. □

CLAIM 19. *For each sequence* $(t_i) \subseteq [0,1)$, *each* $k \in \mathbb{N}^*$ *and any choice of* $\delta_i \in \{0,1\}$, $1 \le i \le k$, *we have*

$$\left|\left\{0 \le t < 1 : \tau_{t_i} r_i(t) = \delta_i, 1 \le i \le k\right\}\right| = \frac{1}{2^k}.$$

Proof. Let $I = \{1 \le i \le k : \delta_i = 1\}$ and $J = \{1 \le i \le k : \delta_i = 0\}$. Writing $E_i = \{0 \le t < 1 : \tau_{t_i} r_i(t) = 1\}$ for each $1 \le i \le k$, we have, according to Claim 18:

$$\begin{aligned}
\left|\left\{0 \le t < 1 : \tau_{t_i} r_i(t) = \delta_i, 1 \le i \le k\right\}\right| &= \int_0^1 \prod_{i\in I} \mathbb{1}_{E_i} \prod_{j\in J} (1 - \mathbb{1}_{E_j}) \\
&= \sum_{J' \subseteq J} (-1)^{\#J'} \int_0^1 \prod_{i\in I} 1_{E_i} \prod_{j\in J'} 1_{E_j} \\
&= \sum_{J' \subseteq J} (-1)^{\#J'} \left(\frac{1}{2}\right)^{\#I} \left(\frac{1}{2}\right)^{\#J'} \\
&= \prod_{i\in I} \frac{1}{2} \prod_{j\in J} \left(1 - \frac{1}{2}\right) \\
&= \prod_{i\in I} \frac{1}{2} \prod_{j\in J} \frac{1}{2} \\
&= \frac{1}{2^k},
\end{aligned}$$

which proves the claim. □

We can now finish the proof of Lemma 17. Fix a sequence $(t_i) \subseteq [0,1)$ and observe that, because the Rademacher functions are characteristic functions on sets with measure $1/2$, it is easy to see that for any Borel sets $E_1, \dots, E_n$, we have

$$\left|\bigcap_{i=1}^n (\tau_{t_i} r_i)^{-1}(E_i)\right| = \prod_{i=1}^n \left|(\tau_{t_i} r_i)^{-1}(E_i)\right|.$$

Actually, the calculation above is proving this when each E_i contains either 0 or 1, and this calculation also implies this result when the E_i either contain just 0, just 1, or both values. The case where some E_i does not contain either 0 or 1 is trivial. This finishes the argument for independence. □

Let us now look at the moving averages that come out of the method of Stokolos. For computational convenience, we are going to work in $\mathbb{T}^2$ where $\mathbb{T} = [0,1] \mod 1$; that is, in the torus.

LEMMA 20. *We assume that that we have pairwise incomparable dyadic rectangles based at the origin $Q_i \subseteq \mathbb{T}^2 \cap [0,1)^2$, $1 \le i \le k$, together with vectors $v_i \in \mathbb{T}^2$, $1 \le i \le k$. We also let $R_i = v_i + Q_i$, $1 \le i \le k$. There are Lebesgue measurable sets Θ and Z in $\mathbb{T}^2$ with the following properties*:

(i) $|Z| \ge \frac{1}{250} k 2^k |\Theta|$;

(ii) *for all $x \in Z$, there exists $\nu_0 = \nu_0(x)$ such that*

$$\frac{|(x - R_{\nu_0}) \cap \Theta|}{|R_{\nu_0}|} \ge \frac{1}{2^{k+1}}.$$

Proof. Write as in Lemma 4 $Q_i = [0, 2^{-m_i}] \times [0, 2^{-n_i}]$ where $1 \le n_1 < \cdots < n_k$ and $m_1 > \cdots > m_k \ge 1$. Define Θ and Y_ν, $1 \le \nu \le k$ as in Lemma 4. For each $1 \le \nu \le k$, let $Y'_\nu = v_\nu + Y_\nu$. We begin by estimating the measure of some relevant intersections.

CLAIM 21. *For each $w \in \mathbb{T}^2$ and any $1 \le \nu, p \le k$, we have $|Y_\nu \cap (w + Y_p)| \le 2^{p-\nu} 2^{-k-1}$.*

Proof. To show this, we write $w = (\alpha, \beta)$ for some $\alpha, \beta \in [0,1)$ and we compute

$$\begin{aligned}
&\big|Y_\nu \cap (w + Y_p)\big| \\
&\quad = \int\!\!\int_{\mathbb{T}^2} \mathbb{1}_{Y_\nu} \mathbb{1}_{w+Y_p} \\
&\quad = \int\!\!\int_{\mathbb{T}^2} \prod_{i_1=\nu}^{k} r_{m_{i_1}}(\xi) \prod_{j_1=1}^{\nu} r_{n_{j_1}}(\eta) \prod_{i_2=p}^{k} r_{m_{i_2}}(\xi - \alpha) \prod_{j_2=1}^{p} r_{n_{j_2}}(\eta - \beta)\, d\xi\, d\eta \\
&\quad \le \int\!\!\int_{\mathbb{T}^2} \prod_{i_1=p}^{s-1} r_{m_{i_1}}(\xi - \alpha) \prod_{i_2=s}^{k} r_{m_{i_2}}(\xi) \prod_{j=1}^{s} r_{n_j}(y)\, d\xi\, d\eta \\
&\quad = \frac{1}{2^{\nu-p}} \frac{1}{2^{k+1}},
\end{aligned}$$

using Lemma 17. □

Using Claim 21, we can now estimate the measure of the set $Y' = \bigcup_{\nu=1}^{k} Y'_\nu$. To that purpose, assume that $k \ge 2$ and denote by $K \ge 1$ the largest integer for which $2K \le k$. We then use the inclusion–exclusion principle to compute

$$\big|Y'\big| \ge \left| \bigcup_{\nu=1}^{K} v_{2\nu} + Y_{2\nu} \right| \ge \sum_{\nu=1}^{K} |Y_{2\nu}| - \sum_{\nu=1}^{K} \sum_{p=1}^{\nu-1} \big|(v_{2\nu} + Y_{2\nu}) \cap (v_{2p} + Y_{2p})\big|.$$

Using the equality $|Y_{2\nu}| = 2^{-k-1}$ valid for each $1 \le \nu \le K$ together with the fact that for $1 \le \nu \le K$ and $1 \le p \le \nu - 1$ we have

$$\big|(v_{2\nu} + Y_{2\nu}) \cap (v_{2p} + Y_{2p})\big| = \big|Y_{2\nu} \cap \big[(v_{2p} - v_{2\nu}) + Y_{2p}\big]\big| \le \frac{1}{4^{\nu-p}} \frac{1}{2^{k+1}}.$$

This follows from Claim 21. Hence,

$$\left|Y'\right| \geq \frac{K}{2^{k+1}} - \frac{1}{2^{k+1}} \sum_{\nu=1}^{K} \sum_{p=1}^{\nu-1} \frac{1}{4^{\nu-p}} = \frac{1}{2^{k+1}} \left[K - \sum_{\nu=1}^{K} \frac{1}{3} \frac{4^{\nu-1}-1}{4^{\nu-1}} \right] \geq \frac{2K}{3} \frac{1}{2^{k+1}}.$$

It follows from the inequality $k \geq 2$ that $K \geq k/4$, so we finally get

$$\left|Y'\right| \geq \frac{k}{6} \frac{1}{2^{k+1}} = \frac{1}{12} k 2^k |\Theta|,$$

for $|\Theta| = 2^{-2k}$.

We now proceed to construct a set Z having the stated properties. Recall that for each $1 \leq \nu \leq k$, Y_ν can be written as

$$Y_\nu = \bigcup_{q=1}^{N_\nu} (u_{\nu,q} + Q_\nu),$$

for some $u_{\nu,q} \in \mathbb{T}^2$, $1 \leq q \leq N_\nu$ where $N_\nu = 2^{-(k+1)} 2^{m_\nu} 2^{n_\nu}$. We then let, for $1 \leq \nu \leq k$,

$$Z_\nu^0 = \bigsqcup_{q=1}^{N_\nu} \left(u_{\nu,q} + Q_\nu^+\right) \quad \text{and} \quad Z_\nu = v_\nu + Z_\nu^0,$$

where we introduced the following notation: given a rectangle $Q = [0,a] \times [0,b]$ with $0 \leq a, b < 1$, we let $Q^+ = [a/2, a] \times [b/2, b]$ denote its upper right corner—similarly we'll also set $Q_- = [0, a/2] \times [0, b/2]$ be its lower left corner. We furthermore let $Z = \bigcup_{\nu=1}^{k} Z_\nu$.

To estimate $|Z|$, we again use the inclusion–exclusion principle as follows: we first assume that $k \geq 4$ and choose $K \geq 1$ the largest integer for which $3K \leq k$. We then have

$$|Z| \geq \left| \bigcup_{\nu=1}^{K} Z_{3\nu} \right| \geq \sum_{\nu=1}^{K} |Z_{3\nu}| - \sum_{\nu=1}^{K} \sum_{p=1}^{\nu-1} \left| \left(v_{3\nu} + Z_{3\nu}^0\right) \cap \left(v_{3p} + Z_{3p}^0\right) \right|.$$

Using the equalities $|Z_{3\nu}^0| = \frac{1}{4} |Y_{3\nu}|$ valid for each $1 \leq \nu \leq K$, together with the inclusions

$$(v_{3\nu} + Z_{3\nu}) \cap (v_{3p} + Z_{3p}) \subseteq (v_{3\nu} + Y_{3\nu}) \cap (v_{3p} + Y_{3p}),$$

valid for each $1 \leq \nu \leq K$ and each $1 \leq p \leq \nu - 1$, we get, using Claim 21 and proceeding as before:

$$|Z| \geq \frac{K}{2^{k+3}} - \sum_{\nu=1}^{K} \sum_{p=1}^{\nu-1} \frac{1}{8^{\nu-p}} \frac{1}{2^{k+1}} = \frac{1}{2^{k+1}} \left[\frac{K}{4} - \sum_{\nu=1}^{K} \frac{1}{7} \frac{8^{\nu-1}-1}{8^{\nu-1}} \right] \geq \frac{3K}{28} \frac{1}{2^{k+1}}.$$

Yet as $k \geq 4$ implies that $K \geq k/12$, we finally obtain:

$$|Z| \geq \frac{k}{125} \frac{1}{2^{k+1}} = \frac{1}{250} k 2^k |\Theta|.$$

Now fix $x \in Z$ and denote by $\nu_{00}(x)$ the smallest integer $1 \le \nu_0 \le k$ for which one has $x \in Z_{\nu_0}$ and observe that

$$\big|(x - R_{\nu_0}) \cap \Theta\big| = \big|\big[x - (v_{\nu_0} + Q_{\nu_0})\big] \cap \Theta\big| = \big|\big[(x - v_{\nu_0}) - Q_{\nu_0}\big] \cap \Theta\big|;$$

noticing that $x - v_{\nu_0}$ belongs to $u_{\nu_0,q} + Q^+_{\nu_0}$ for some $1 \le q \le N_{\nu_0}$, we see that $u_{\nu_0,q} + Q^-_{\nu_0} \subseteq (x - v_{\nu_0}) - Q_{\nu_0}$ (to see this, say that $x - v_{\nu_0} = u_{\nu_0,q} + y$, $y \in Q^+_{\nu_0}$; now if we take $z \in Q^-_{\nu_0}$, we have to show that

$$u_{\nu_0,q} + z \in u_{\nu_0,q} + y - Q_{\nu_0};$$

this is obvious for $y - z \in Q_{\nu_0}$). We hence have

$$\big|(x - R_{\nu_0}) \cap \Theta\big| \ge \big|\big(u_{\nu_0,q} + Q^-_{\nu_0}\big) \cap \Theta\big| = \big|(u_{\nu_0,q} + Q_{\nu_0}) \cap \Theta_{\nu_0}\big|,$$

where $\Theta_{\nu_0} = \{(\xi,\eta) \in \Theta : r_{m_{\nu_0}-1}(\xi) r_{n_{\nu_0}+1}(\eta) = 1\} \subseteq Y_{\nu_0}$. It is obvious that $|\Theta_{\nu_0}| \ge \frac{1}{4}|\Theta|$. On the other hand, as in Lemma 4, we now observe that Y_{ν_0} is an union of N_{ν_0} rectangles $u_{\nu_0,q} + Q_{\nu_0}$, $1 \le q \le N_{\nu_0}$ and that for any $1 \le q, r \le N_{\nu_0}$ we have

$$\big|(u_{\nu_0,q} + Q_{\nu_0}) \cap \Theta_{\nu_0}\big| = \big|(u_{\nu_0,r} + Q_{\nu_0}) \cap \Theta_{\nu_0}\big|.$$

We then infer from the inclusion $\Theta_{\nu_0} \subseteq Y_{\nu_0}$ that

$$\frac{|(u_{\nu_0,q} + Q_{\nu_0}) \cap \Theta_{\nu_0}|}{|u_{\nu_0,q} + Q_{\nu_0}|} = \frac{N_{\nu_0}|(u_{\nu_0,q} + Q_{\nu_0}) \cap \Theta_{\nu_0}|}{N_{\nu_0}|u_{\nu_0,q} + Q_{\nu_0}|} = \frac{|\Theta_{\nu_0}|}{|Y_{\nu_0}|} \ge \frac{1}{4}\frac{|\Theta|}{|Y_{\nu_0}|} = \frac{1}{2^{k+1}}.$$

The proof is complete for $|u_{\nu_0,q} + Q_{\nu_0}| = |R_{\nu_0}|$. □

Given a sequence $\mathbf{R} = (R_i)$ of rectangles, we define a maximal operator $M_{\mathbf{R}}$ on $L^1(\mathbb{R}^2)$ by

$$M_{\mathbf{R}} f(x) = \sup_{i \in \mathbb{N}^*} \frac{1}{|R_i|} \mathbb{1}_{R_i} * f(x). \tag{5}$$

Lemma 20 is a key tool in the study of the behaviour of the maximal operator $M_{\mathbf{R}}$ when the sequence $\mathbf{R}$ is obtained from a sequence (Q_i) of standard rectangles in $[0,1)^2$ by translations:

$$R_i = v_i + Q_i \quad \text{where } v_i \in \mathbb{R}^2, i \in \mathbb{N}^*.$$

It allows us to prove an analogue of Corollary 8 in the moving context.

COROLLARY 22. *Let* $\mathbf{Q}$ *be a family of standard rectangles in* $[0,1)^2$, *fix a sequence* $(v_i) \subseteq \mathbb{R}^2$, *let* $\mathbf{R} = (R_i)$ *be the sequence of rectangles in* $\mathbb{R}^2$ *defined by* $R_i = v_i + Q_i$ *and assume that we have* $R_i \subseteq [0,1)^2$ *for each* $i \in \mathbb{N}^*$. *If* $\mathbf{Q}^*$ *has infinite width, then for each* $\lambda \ge 1$ *there exists a sequence of functions* $(f_k) \subseteq L^\infty_+(\mathbb{R}^2)$ *satisfying the following conditions*: *for each* $k \in \mathbb{N}^*$,

(i) f_k *vanishes outside* $[0,1]^2$;
(ii) $\int_{\mathbb{R}^2} \Phi_0(|f_k|)\,dx \ge \frac{1}{2}(k+1)\|f_k\|_1$;
(iii) $|\{x \in \mathbb{R}^2 : M_{\mathbf{R}} f_k(x) \ge \lambda\}| \ge \frac{1}{500\lambda} \int_{\mathbb{R}^2} \Phi_0(|f_k|)\,dx$;

(iv) *for any Orlicz function* Φ *satisfying* $\Phi = o(\Phi_0)$ *at* ∞ *and for each* $C > 0$ *we have*

$$\lim_{k\to\infty} \frac{\int_{\mathbb{R}^2} \Phi_0(|f_k|)\,dx}{\int_{\mathbb{R}^2} \Phi(C|f_k|)\,dx} = \infty.$$

Proof. We view $[0,1)^2$ as a subset of $\mathbb{T}^2$. Take $k \in \mathbb{N}^*$. As $\mathbf{Q}$ has infinite width, find indices $1 \le i_1 < i_2 < \cdots < i_k$ such that $\{Q_{i_j} : 1 \le j \le k\}$ is independent. Choose sets Θ_k and Z_k in $\mathbb{T}^2$ associated to Q_{i_j} and v_{i_j}, $1 \le i \le k$ according to Lemma 20. Let $f_k = 2^{k+1}\lambda \mathbb{1}_{\Theta_k}$. It is clear that f_k is supported in $[0,1]^2$, and (i) is proved.

To prove (ii), begin by observing that

$$\begin{aligned}\int_{\mathbb{R}^2} \Phi_0\big(|f_k|\big)\,dx &= |\Theta_k| 2^{k+1}\lambda\big[1 + \log\lambda + (k+1)\log 2\big] \\ &= 2^{1-k}\lambda\big[1 + \log\lambda + (k+1)\log 2\big].\end{aligned}$$

On the other hand, we have

$$\|f_k\|_1 = 2^{k+1}\lambda|\Theta_k| = 2^{1-k}\lambda.$$

We hence get:

$$\int_{\mathbb{R}^2} \Phi_0\big(|f_k|\big)\,dx = \big[1 + \log\lambda + (k+1)\log 2\big]\|f_k\|_1 \ge \frac{1}{2}(k+1)\|f_k\|_1.$$

We prove (iii) as follows: fix $x \in Z_k$ and choose an integer $1 \le \nu_0 = \nu_0(x) \le k$ according to Lemma 20(ii). Compute now

$$M_{\mathbf{R}} f_k(x) \ge \frac{1}{|R_{i_{\nu_0}}|}\mathbb{1}_{R_{i_{\nu_0}}} * f_k(x) = 2^{k+1}\lambda \frac{|(x - R_{i_{\nu_0}}) \cap \Theta_k|}{|R_{i_{\nu_0}}|} \ge \lambda.$$

We hence have $Z_k \subseteq \{x \in \mathbb{R}^2 : M_{\mathbf{R}} f_k(x) \ge \lambda\}$. It follows that for k sufficiently large, we have

$$\begin{aligned}\big|\big\{x \in \mathbb{R}^2 : M_{\mathbf{R}} f_k(x) \ge \lambda\big\}\big| \ge |Z_k| &\ge \frac{1}{250} k 2^k |\Theta_k| = \frac{1}{500\lambda} k\big(2^{1-k}\lambda\big) \\ &= \frac{1}{500\lambda} \frac{k}{1 + \log\lambda + (k+1)\log 2} \int_{\mathbb{R}^2} \Phi_0\big(|f_k|\big)\,dx \\ &\ge \frac{1}{500\lambda} \int_{\mathbb{R}^2} \Phi_0\big(|f_k|\big)\,dx.\end{aligned}$$

Finally, the proof of (iv) is virtually identical to the proof of Corollary 8(iv). □

REMARK 23. Under the hypotheses of Corollary 22, it now follows immediately from Corollary 22 that the maximal operator $M_{\mathbf{R}}$ cannot satisfy a weak

$(1,1)$ inequality, for otherwise there would exist a constant $C>0$ independent of k such that for any sufficiently large $k\in\mathbb{N}^*$ we have

$$\begin{aligned}\frac{1}{500}\int_{\mathbb{R}^2}\Phi_0\big(|f_k|\big)\,dx &\le \big|\big\{x\in\mathbb{R}^2:\,M_{\mathbf{R}}f_k(x)\ge 1\big\}\big|\\ &\le \big|\big\{x\in\mathbb{R}^2:\,M_{\mathbf{R}}f(x)>1/2\big\}\big|\le 2C\|f_k\|_1,\end{aligned}$$

which would contradict Corollary 22(ii) for k sufficiently large.

Remark 24. Under the same hypotheses, it also follows from Corollary 22(iv) that the maximal operator $M_{\mathbf{R}}$ cannot satisfy an inequality of weak type in $L^{\Phi}(\mathbb{R}^2)$, in case Φ is a Young function satisfying $\Phi=o(\Phi_0)$ at ∞, where Φ_0 is the Orlicz function in Example 41. To see this, we observe that if there would exist a constant $C>0$ such that

$$\big|\big\{x\in\mathbb{R}^2:\,M_{\mathbf{R}}f(x)>\lambda\big\}\big|\le\int_{\mathbb{R}^2}\Phi\bigg(\frac{C|f_k|}{\lambda}\bigg)\,dx,$$

holds for every $f\in L^{\Phi}_+(\mathbb{R}^2)$ and every $\lambda>0$, then if (f_k) is the sequence of functions coming out of Corollary 22, we would have for k sufficiently large:

$$\begin{aligned}\frac{1}{500}\int_{\mathbb{R}^2}\Phi_0\big(|f_k|\big)\,dx &\le \big|\big\{x\in\mathbb{R}^2:\,M_{\mathbf{R}}f_k(x)\ge 1\big\}\big|\\ &\le \bigg|\bigg\{x\in\mathbb{R}^2:\,M_{\mathbf{R}}f_k(x)>\frac{1}{2}\bigg\}\bigg|\le\int_{\mathbb{R}^2}\Phi\big(2C\|f_k\|\big)\,dx,\end{aligned}$$

which would contradict Corollary 22(iv) for k sufficiently large.

We summarize the facts about the behaviour of $M_{\mathbf{R}}$ that come out from the preceding remarks, in the following result.

Proposition 25. *Assume that the hypotheses of Corollary 22 are satisfied. If Φ is an Orlicz function satisfying $\Phi=o(\Phi_0)$ at ∞, where Φ_0 is the Orlicz function of Example 41, then there exists $f\in L^{\Phi}_+(\mathbb{R}^2)$ for which $M_{\mathbf{R}}f(x)=\infty$ holds a.e. in $\mathbb{R}^2$; in particular, there exists $f\in L^1_+(\mathbb{R}^2)$ such that $M_{\mathbf{R}}f=\infty$ a.e. on $\mathbb{R}^2$.*

Proof. The proof is virtually identical to the proof of Theorem 10(ii). □

Remark 26. As a counterpart to Proposition 25, we refer to the end of Section 2.2 for a *positive* result along the lines of Theorem 10(i).

2. Moving averages in the plane: The ergodic context

In this section, we fix a Lebesgue probability space (X,μ) together with commuting, invertible measure-preserving transformations $S,T:X\to X$. We moreover assume that the action

$$\mathbb{Z}^2\times X\to X,\qquad (k,l)\mapsto S^kT^lx$$

is *free*, i.e. that $\mu\{x:\,S^kT^lx=x\}=0$ unless $k=l=0$.

2.1. Standard averages in the ergodic context: The L^1 behaviour. A *standard rectangle* in $\mathbb{Z}^2$ is a set $Q \subseteq \mathbb{R}^2$ of the form $Q = \mathbb{Z}^2 \cap [0,m] \times [0,n]$ for integers $m, n \in \mathbb{N}^*$; we then let $\ell(Q) = m$, $L(Q) = n$, while $\bar{Q}$ will stand for the associated rectangle $\bar{Q} = [0,m] \times [0,n]$ in $\mathbb{R}^2$. Conversely, given a standard rectangle Q in $\mathbb{Z}^2$, we let $\lfloor Q \rfloor$ denote the largest standard rectangle in $\mathbb{Z}^2$ contained in Q.

A *standard dyadic rectangle* in $\mathbb{Z}^2$ is a rectangle of the form $[0, 2^m] \times [0, 2^n]$. The *dyadic mother* Q^* of a standard rectangle Q in $\mathbb{Z}^2$ is the standard dyadic rectangle containing Q that has the least number of elements.

For our purposes, we will call an *admissible sequence* of standard rectangles any sequence $\mathbf{Q} = (Q_i)$ of standard rectangles in $\mathbb{Z}^2$ such that both sequences $(\ell(Q_i))$ and $(L(Q_i))$ tend to ∞.

The definitions we made in the differentiation context naturally generalize to the present setting.

DEFINITION 27. Two standard rectangles Q_1 and Q_2 in $\mathbb{Z}^2$ are called *incomparable*, and we write $Q_1 \nsim Q_2$, in case neither $Q_1 \subseteq Q_2$ nor $Q_2 \subseteq Q_1$.

Moreover, a family $\mathbf{Q}$ of standard rectangles in $\mathbb{Z}^2$ is called

- *independent* in case $Q_1 \nsim Q_2$ holds for any distinct $Q_1, Q_2 \in \mathbf{Q}$;
- *dependent* in case there exist distinct elements $Q_1, Q_2 \in \mathbf{Q}$ where $Q_1 \nsim Q_2$ does not hold;
- *a chain* in case $\mathbf{Q}$ is totally ordered by inclusion.

DEFINITION 28. A sequence $\mathbf{Q}$ of standard rectangles in $\mathbb{Z}^2$ is said to have *infinite width* in case for every $k \in \mathbb{N}$, there exists integers $1 \leq i_1 < i_2 < \cdots < i_k$ for which $\{Q_{i_j} : 1 \leq j \leq k\}$ is independent. It is said to have *finite width* in case it is not of infinite width.

Dilworth's alternative still holds.

LEMMA 29. *A sequence $\mathbf{Q}$ of standard rectangles in $\mathbb{Z}^2$ has either infinite width, and it is not a finite union of chains, or it is finite width and it is a finite union of chains.*

To any admissible sequence $\mathbf{Q} = (Q_i)$ of standard rectangles in $\mathbb{Z}^2$, one naturally associates a maximal operator $M_{\mathbf{Q}}$ defined on $L^1(X, \mu)$ by

$$M_{\mathbf{Q}} f(x) = \sup_{i \in \mathbb{N}^*} \frac{1}{\#Q_i} \sum_{(k,l) \in Q_i} f\big(S^k T^l x\big).$$

The behavior of $M_{\mathbf{Q}}$ may be studied according to the comparability properties of the dyadic approximations of its elements.

THEOREM 30. *Let $\mathbf{Q}$ be a sequence of standard rectangles in $\mathbb{Z}^2$. The following properties are satisfied*:

(i) *if $\mathbf{Q}^*$ has finite width, then for any $f \in L^1(X, \mu)$ we have $M_{\mathbf{Q}} f < \infty$, μ-a.e. on X;*

(ii) *if* $\mathbf{Q}^*$ *has infinite width and if* Φ *is an Orlicz function satisfying* $\Phi = o(\Phi_0)$ *at* ∞, *then there exists* $f \in L^{\Phi}_{+}(X,\mu)$ *such that* $M_{\mathbf{Q}} f = \infty$ *holds* μ*-a.e. on* X; *in particular then there exists* $f \in L^1_+(X,\mu)$ *for which* $M_{\mathbf{Q}} f = \infty$, μ*-a.e. on* X.

Proof. To prove (i), assume that $\mathbf{Q}^*$ has finite width; it then follows from Hagelstein and Stokolos [7, Theorem 1] that the maximal operator $M_{\mathbf{Q}}$ is of weak type $(1,1)$. One hence finishes the proof of (i) by applying Sawyer [13, Theorem 3].

As the proof of (ii) would need a more general transfer lemma along the lines of Lemma 42 in order to transfer the inequality (2) in the ergodic context, we only prove (ii) in $L^1(\mathbb{R}^2)$ for brevity's sake. To this purpose let us proceed towards a contradiction, and assume that (ii) does not hold. In this case, it follows from Stein [14, Corollary 1] that the maximal operator $M_{\mathbf{Q}}$ satisfies an inequality of weak type $(1,1)$. By Rokhlin's lemma as in Ornstein and Weiss [10], this implies that the maximal operator $m_{\mathbf{Q}}$ defined on $\ell^1(\mathbb{Z}^2)$ by

$$m_{\mathbf{Q}}\varphi(m,n) = \sup_{i\in\mathbb{N}^*} \frac{1}{\#Q_i} \sum_{(k,l)\in Q_i} \big|\varphi(m+k,n+l)\big|$$

for $m,n \in \mathbb{Z}$, satisfies an inequality of weak type $(1,1)$ in $\ell^1(\mathbb{Z}^2)$.

By our transfer lemma (Lemma 42, see Appendix B), there would exist a constant $C>0$ such that for any $f \in L^1(\mathbb{R}^2)$ and any $\lambda > 0$:

$$\big|\big\{x \in \mathbb{R}^2 : M_{\bar{\mathbf{Q}}} f(x) > \lambda\big\}\big| \le \frac{C\|f\|_1}{\lambda};$$

where $\bar{\mathbf{Q}} = (\bar{Q}_i)$ is the sequence of standard rectangles in $\mathbb{R}^2$ associated to Q_i, $i \in \mathbb{N}^*$.

On the other hand, observe now that if we define a sequence $\tilde{\mathbf{Q}} = (\tilde{Q}_i)$ of standard rectangles in $\mathbb{Z}^2$ by $\tilde{Q}_i = \lfloor \bar{Q}_i/2 \rfloor$, then $\tilde{\mathbf{Q}}^* = (\tilde{Q}_i^*)$ also has infinite width. For each $i \in \mathbb{N}^*$, let $Q_i' = \tilde{Q}_i^*$ and observe that (using the notations of Lemma 42) we have $f_{\mathbb{1}_{Q_i'}} = \mathbb{1}_{\bar{Q}_i'}$. Fix $k \in \mathbb{N}^*$. According to the fact that $\tilde{\mathbf{Q}}^*$ has infinite width, find integers $1 \le i_1 < i_2 < \cdots < i_k$ such that $\bar{Q}_{i_1}', \ldots, \bar{Q}_{i_k}'$ form an independent family of standard rectangles in $\mathbb{R}^2$. Choose next an $a>0$ such that we have $\bar{Q}_{i_j}' \subseteq [0,a)^2$ for each $1 \le j \le k$, let $Q_j'' = a^{-1}\bar{Q}_{i_j}'$ for each $1 \le j \le k$ and let f_k denote the function associated to $\bar{Q}_1'', \ldots, \bar{Q}_k''$ and $\lambda = 5$ by Corollary 8.

For each $k \in \mathbb{N}^*$, define $g_k = d_a f_k$, where the dilation $d_a f_k \in L^1(\mathbb{R}^2)$ is defined at $x \in \mathbb{R}^2$ by the formula $d_a f_k(x) = a^2 f_k(ax)$. Observe in particular that, for any $x \in \mathbb{R}^2$ and $1 \le j \le k$, we have

$$\mathbb{1}_{Q_j''} * f_k(x) = \int_{Q_j''} f_k(x-\xi)\,d\xi = \int_{\bar{Q}_{i_j}'} a^2 f_k(x-a\eta)\,d\eta = \mathbb{1}_{\bar{Q}_j'} * g_k(x/a).$$

It follows from the proof of Corollary 8(ii), (iii) that we have, for k sufficiently large,

$$\left|\left\{x\in\mathbb{R}^2:\max_{1\le j\le k}\frac{1}{|Q_j''|}\mathbb{1}_{Q_j''}*f_k(x)\ge 5\right\}\right|\ge\frac{1}{20}k\|f_k\|_1.$$

Using the equality $|\bar{Q}'_{i_j}|=a^2|Q''_j|$, we hence get, for k sufficiently large:

$$\begin{aligned}&\left|\left\{x\in\mathbb{R}^2:\max_{1\le j\le k}\frac{1}{|\bar{Q}'_{i_j}|}\mathbb{1}_{\bar{Q}'_{i_j}}*g_k(x)\ge 5a^2\right\}\right|\\&\quad=\frac{1}{a^2}\left|\left\{y\in\mathbb{R}^2:\max_{1\le j\le k}\frac{1}{|Q''_j|}\mathbb{1}_{Q''_j}*f_k(y)\ge 5\right\}\right|\\&\quad\ge\frac{1}{20a^2}k\|f_k\|_1=\frac{1}{20a^2}k\|g_k\|_1.\end{aligned}$$

We hence get, using the inequality $M_{\mathbf{Q}}g_k(x)\ge\frac{1}{4}M_{\tilde{\mathbf{Q}}^*}g_k(x)$, valid for each $x\in\mathbb{R}^2$ and each $k\in\mathbb{N}^*$ sufficiently large:

$$\begin{aligned}\frac{C\|g_k\|_1}{a^2}&\ge\big\|\big\{x\in\mathbb{R}^2:M_{\mathbf{Q}}g_k(x)>a^2\big\}\big\|\\&\ge\left|\left\{x\in\mathbb{R}^2:\max_{1\le j\le k}\frac{1}{|\bar{Q}'_{i_j}|}\mathbb{1}_{\bar{Q}'_{i_j}}*g_k(x)\ge 5a^2\right\}\right|\ge\frac{1}{20a^2}k\|g_k\|_1;\end{aligned}$$

which is impossible when k is large enough. □

REMARK 31. We can also prove an analogue of Theorem 10(iii), but do not do so here for brevity's sake.

REMARK 32. It is noteworthy to observe that an analogue of Remark 16 can be stated in the ergodic context, for admissible sequences $\mathbf{Q}=(Q_i)$ of standard rectangles in $\mathbb{Z}^2$.

2.2. Moving averages in the ergodic context: The L^1 behavior. We first recall a result by Bellow, Jones and Rosenblatt in [2] concerning one-dimensional moving averages. The context is the following: let $\Omega\subseteq\mathbb{Z}\times\mathbb{N}^*$ and define, for each $\alpha>0$:

$$\Omega_\alpha=\big\{(k,r)\in\Omega:\big|k-k'\big|\le\alpha\big(r-r'\big)\text{ for some }\big(k',r'\big)\in\Omega\big\}.$$

One further defines, given $r\in\mathbb{N}^*$ and $\alpha>0$, the cross-section

$$\Omega_\alpha(r)=\big\{k\in\mathbb{Z}:(k,r)\in\Omega_\alpha\big\}.$$

One also associates to Ω a maximal operator M_Ω defined on $L^1(X,\mu)$ by

$$M_\Omega f(x)=\sup_{(k,r)\in\Omega}\frac{1}{r+1}\sum_{j=0}^{r}\big|f\big(T^{k+j}x\big)\big|.$$

THEOREM 33 (Bellow, Jones and Rosenblatt). *The following two assertions are satisfied*:

(i) *if there exist constants* $\alpha > 0$ *and* $A \geq 0$ *having the property that for any* $r \in \mathbb{N}^*$, $\#\Omega_\alpha(r) \leq Ar$, *then* M_Ω *is of weak type* $(1,1)$ *and of strong type* (p,p) *for any* $1 < p \leq \infty$.

(ii) *if* M_Ω *is of weak type* (p,p) *for some* $1 < p < \infty$, *then for any* $\alpha > 0$ *there exists* $A_\alpha \geq 0$ *having the property that for any* $r \in \mathbb{N}^*$, *then* $\#\Omega_\alpha(r) \leq A_\alpha r$.

We now let $\mathbf{Q} = (Q_i)$ be a sequence of standard rectangles in $\mathbb{Z}^2$, and we fix a sequence $\mathbf{v} = (v_i) \subseteq \mathbb{N}^2$; we also define a sequence $\mathbf{R} = (R_i)$ of rectangles in $\mathbb{Z}^2$ by letting $R_i = v_i + Q_i$ for each $i \in \mathbb{N}$. The maximal operator $M_{\mathbf{R}}$ is defined in the natural way by

$$M_{\mathbf{R}} f(x) = \sup_{i \in \mathbb{N}} \frac{1}{\#R_i} \sum_{(k,l) \in R_i} f\big(S^k T^l x\big).$$

Denoting by p and q the orthogonal projections on the x- and y-axis, respectively, we associate to the sequence $\mathbf{R}$, the following nets:

$$\Omega(\mathbf{v}, \mathbf{Q}) = \big\{ \big(p(v_i), \#p(Q_i)\big) : i \in \mathbb{N} \big\} \subseteq \mathbb{N} \times \mathbb{N}^*,$$
$$\mho(\mathbf{v}, \mathbf{Q}) = \big\{ \big(q(v_i), \#q(Q_i)\big) : i \in \mathbb{N} \big\} \subseteq \mathbb{N} \times \mathbb{N}^*.$$

LEMMA 34. *Assume that* $\mathbf{Q}$ *is a sequence of standard rectangles in* $\mathbb{Z}^2$, *and fix a sequence* $\mathbf{v} = (v_i) \subseteq \mathbb{N}^2$. *As usual, define a sequence of rectangles* $\mathbf{R} = (R_i)$ *by letting* $R_i = v_i + Q_i$ *for* $i \in \mathbb{N}$. *Consider the following two statements*:

(i) *the maximal operator* $M_{\mathbf{R}}$ *is of weak type* $(1,1)$;

(ii) *there exists constants* $\alpha > 0$ *and* $A \geq 0$ *such that for any* $r \in \mathbb{N}^*$, *we have*

$$\#\big(\Omega(\mathbf{v}, \mathbf{Q})\big)_\alpha(r) \leq Ar \quad \text{and} \quad \#\big(\mho(\mathbf{v}, \mathbf{Q})\big)_\alpha(r) \leq Ar.$$

Then, (i) *implies* (ii) *and the converse holds* in case $\mathbf{Q}$ is a chain of rectangles.

Proof. To show that (i) implies (ii), begin by observing that Rokhlin's lemma (see Ornstein and Weiss [10]) together with Calderón's transfer principle (see Calderón [4]) implies that for *any* commuting, measure-preserving transformations $S', T' : X \to X$, the maximal operator $M'_{\mathbf{R}}$ defined on $L^1(X, \mu)$ by

$$M'_{\mathbf{R}} f(x) = \sup_{i \in \mathbb{N}} \frac{1}{\#R_i} \sum_{(k,l) \in R_i} f\big(S'^k T'^l x\big),$$

is of weak type $(1,1)$ provided (i) holds. Taking either $S' = S$ and $T' = \mathrm{id}_X$ or $S' = \mathrm{id}_X$ and $T' = T$, and observing that for any $i \in \mathbb{N}$:

$$\frac{1}{\#p(Q_i)} \sum_{j=1}^{\#p(Q_i)} f\big(S^{p(v_i)+j} x\big) = \frac{1}{\#R_i} \sum_{(k,l) \in R_i} f\big(S^k \,\mathrm{id}_X^l\, x\big),$$

$$\frac{1}{\#q(Q_i)} \sum_{j=1}^{\#q(Q_i)} f\big(T^{q(v_i)+j} x\big) = \frac{1}{\#R_i} \sum_{(k,l) \in R_i} f\big(\mathrm{id}_X^k\, T^l x\big),$$

we see that the maximal operators $M_{\Omega(\mathbf{v},\mathbf{Q})}$ and $M_{\mho(\mathbf{v},\mathbf{Q})}$ are of weak type $(1,1)$. Assertion (ii) then follows from Bellow, Jones and Rosenblatt theorem (Theorem 33).

Assume now that $\mathbf{Q}$ is increasing and satisfies (ii). According to Theorem 33 and to Avramidou [1, Theorem 1.1], we obtain that the correct factors (which, in this case, equal Avramidou's *modified correct factors*)

$$M_i(\mathbf{R}) = \#\cup\big\{p(R_i) - p(R_j) : 1 \le j \le i\big\} \quad \text{and}$$
$$N_i(\mathbf{R}) = \#\cup\big\{q(R_i) - q(R_j) : 1 \le j \le i\big\}$$

satisfy the inequalities

$$M_i(\mathbf{R}) \le C\#p(R_i) \quad \text{and} \quad N_i(\mathbf{R}) \le C\#q(R_i)$$

for each $i \in \mathbb{N}$, where $C > 0$ is a constant independent from i. Yet defining the joint correct factor

$$P_i(\mathbf{R}) = \#\cup\{R_i - R_j : 1 \le j \le i\}$$

for $i \in \mathbb{N}$, and observing that we have $P_i(\mathbf{R}) \le M_i(\mathbf{R})N_i(\mathbf{R})$, we see that $P_i(\mathbf{R}) \le C^2\#R_i$ holds for each $i \in \mathbb{N}$; it hence follows from Rosenblatt and Wierdl [12, Theorem 5.8] that $M_{\mathbf{R}}$ is of weak type $(1,1)$. □

We now state the analogue of Theorem 30 in the context of moving averages.

THEOREM 35. *Let* $\mathbf{Q}$, $\mathbf{v}$ *and* $\mathbf{R}$ *be as above, and let* $\Omega = \Omega(\mathbf{v},\mathbf{Q})$ *and* $\mho = \mho(\mathbf{v},\mathbf{Q})$ *be the associated nets. The following properties hold*:

(i) *assume that* $\mathbf{Q}^*$ *has finite width*; *if moreover there exists constants* $\alpha > 0$ *and* $A > 0$ *such that for any* $r \in \mathbb{N}^*$ *we have*:

$$\#\Omega_\alpha(r) \le Ar \quad \text{and} \quad \#\mho_\alpha(r) \le Ar,$$

then for any $f \in L^1(X,\mu)$ *we have* $M_{\mathbf{R}}f < \infty$, μ*-a.e. on* X;

(ii) *if* $\mathbf{Q}^*$ *has infinite width and if* Φ *is an Orlicz function satisfying* $\Phi = o(\Phi_0)$ *at* ∞, *then there exists* $f \in L^\Phi_+(X,\mu)$ *such that* $M_{\mathbf{R}}f = \infty$, μ*-a.e. on* X; *in particular there exists* $f \in L^1_+(X,\mu)$ *for which* $M_{\mathbf{Q}}f = \infty$, μ*-a.e. on* X.

Proof. The proof of (ii) is virtually identical to the proof of Theorem 30, and relies on Corollary 22 instead of Corollary 8 in the case of $L^1(X,\mu)$; for brevity's sake, we omit the proof in the general Orlicz space $L^\Phi(X,\mu)$.

To prove (i), let $\mathbf{Q}$, $\mathbf{v}$ and $\mathbf{R}$ be as above, and assume that $\mathbf{Q}^*$ has finite width. Writing—according to Dilworth's alternative (Lemma 29)—$\mathbf{Q}^* = \mathbf{Q}^*_1 \cup \cdots \cup \mathbf{Q}^*_n$ where the $\mathbf{Q}^*_j$, $1 \le j \le n$ are increasing sequences of rectangles extracted from $\mathbf{Q}^*$, let $(\mathbf{v}_j)$, $1 \le j \le n$ be the corresponding subsequences of $\mathbf{v}$. Condition (i) then easily follows from Sawyer [13, Theorem 1] since the following two statements are equivalent:

(A) for any $f \in L^1(X,\mu)$, we have $M_{\mathbf{R}}f < \infty$ for μ-a.e. on X;

(B) there exists constants $\alpha > 0$ and $A > 0$ such that for any $r \in \mathbb{N}^*$ and any $1 \leq j \leq n$, we have

$$\#\big(\Omega\big(\mathbf{v}_j, \mathbf{Q}_j^*\big)\big)_\alpha(r) \leq Ar \quad \text{and} \quad \#\big(\mathcal{U}\big(\mathbf{v}_j, \mathbf{Q}_j^*\big)\big)_\alpha(r) \leq Ar.$$

It indeed follows from Lemma 34 and from simple computations analogous to those made in the proof of Theorem 10(i). □

REMARK 36. Going back to the differentiation context as announced in Remark 26, and letting $\mathbf{Q}$ be a sequence of standard rectangles in $\mathbb{R}^2$ and $\mathbf{v} = (v_i) \subseteq \mathbb{R}^2$ be a sequence of vectors in $\mathbb{R}^2$, we denote, as expected, by p and q the projections on the x- and y- axes, respectively.

Given a set $\Omega \subseteq \mathbb{R} \times [0, \infty)$, we let, for any $\alpha > 0$:

$$\Omega_\alpha := \big\{(x,t) \in \mathbb{R} \times [0,\infty) : \big|x - x'\big| \leq \alpha \big|t - t'\big| \text{ for some } \big(x', t'\big) \in \Omega\big\},$$

while we let, for $t \geq 0$:

$$\Omega(t) := \big\{x \in \mathbb{R} : (x,t) \in \Omega\big\}.$$

To $\mathbf{Q}$ and $\mathbf{v}$, we then associate two sets

$$\Omega(\mathbf{v}, \mathbf{Q}) = \big\{\big(p(v_i), \big|p(Q_i)\big|\big) : i \in \mathbb{N}\big\} \subseteq \mathbb{R} \times [0,\infty),$$
$$\mathcal{U}(\mathbf{v}, \mathbf{Q}) = \big\{\big(q(v_i), \big|q(Q_i)\big|\big) : i \in \mathbb{N}\big\} \subseteq \mathbb{R} \times [0,\infty).$$

Using Nagel and Stein [9, Section 2, Theorem 1] instead of Bellow, Jones and Rosenblatt (Theorem 33), we could then state an analogue result to [Theorem 35(i) in the context of differentiation:

THEOREM 37. *Assume that $\mathbf{Q}^*$ is a chain of rectangles in $\mathbb{R}^2$, let $\mathbf{R} = (R_i)$ be defined by $R_i := v_i + Q_i$ for $i \in \mathbb{N}$ and let $\Omega := \Omega(\mathbf{v}, \mathbf{Q})$ and $\mathcal{U} := \mathcal{U}(\mathbf{v}, \mathbf{Q})$ be the associated sets. Then the maximal operator $M_{\mathbf{R}}$ defined in (5) is of weak type $(1,1)$ in case there exists constants $\alpha > 0$ and $A \geq 0$ such that for any $0 \leq t < \infty$, we have*

$$\big|\Omega_\alpha(t)\big| \leq At \quad \text{and} \quad \big|\mathcal{U}_\alpha(t)\big| \leq At.$$

The proof of Theorem 35(i) indeed translates to the differentiation context.

2.3. Further results in higher exponent Lebesgue spaces. In this whole section, we assume, as before, that $\mathbf{Q}$ is a sequence of standard rectangles in $\mathbb{Z}^2$, we fix a sequence $\mathbf{v} = (v_i)$ in $\mathbb{N}^2$ and define a sequence $\mathbf{R} = (R_i)$ of rectangles in $\mathbb{Z}^2$ by $R_i = v_i + Q_i$, $i \in \mathbb{N}$. We also define operators A_i and B_i on $L^1(X, \mu)$ by

$$A_i f(x) = \frac{1}{\#p(R_i)} \sum_{k \in p(R_i)} f\big(S^k x\big) \quad \text{and} \tag{6}$$
$$B_i f(x) = \frac{1}{\#q(R_i)} \sum_{l \in q(R_i)} f\big(T^l x\big),$$

and we denote by A_* and B_* the associated maximal operators.

Assuming that the maximal operator $M_{\mathbf{R}}$ is of weak type (p_0,p_0) for some $1<p_0<\infty$, it follows from Rokhlin's lemma (see Ornstein and Weiss [10]) and from the Calderón transfer principle (see Calderón [4]) that for any pair S',T' of commuting, measure preserving transformations of X, the maximal operator $M'_{\mathbf{R}}$ associated in the obvious way to $\mathbf{R}$, S' and T', is of weak type (p_0,p_0). Denoting by (A'_i) and (B'_i) the sequence of linear operators associated in the obvious way to $\mathbf{R}$, S' and T' as in (6), we observe that a computation along the lines of the first part of the proof of Lemma 34 shows that the two associated maximal operators A'_* and B'_* are of weak type (p_0,p_0). Applying Bellow, Jones and Rosenblatt theorem (Theorem 33) *twice* then yields the fact that both maximal operators A_* and B_* are of weak type $(1,1)$ and of strong type (p,p) for any $1<p\leq\infty$. It then follows from Zygmund [17, II, Theorem 4.34 and Remark p. 119] that there exists a constant $C>0$ such that for any $f\in L\log L(X,\mu)$, we have

$$\max\bigl\{\bigl\|A'_*f\bigr\|_1,\bigl\|B'_*f\bigr\|_1\bigr\}\leq C\biggl(1+\int_X\Phi_0\bigl(|f|\bigr)\,d\mu\biggr). \tag{7}$$

On the other hand, one easily observes that for any $i\in\mathbb{N}$ and $f\in L^1_+(X,\mu)$, we have

$$\frac{1}{\#R_i}\sum_{(k,l)\in R_i}f\bigl(S'^kT'^lx\bigr)=A'_i\circ B'_if(x);$$

we thus have, for $i\in\mathbb{N}$:

$$\frac{1}{\#R_i}\sum_{(k,l)\in R_i}f\bigl(S'^kT'^lx\bigr)=A'_i\circ B'_if(x)\leq A'_i\circ B'_*f(x),$$

which yields the inequality

$$M'_{\mathbf{R}}f\leq A'_*\circ B'_*. \tag{8}$$

Yet given $f\in L\log L_+(X,\mu)$, equation (7) shows that we have $B'_*f\in L^1_+(X,\mu)$. According to the fact that A'_* is of weak type $(1,1)$ (see the discussion above) we hence get

$$M'_{\mathbf{R}}f(x)=A'_*\circ B'_*f(x)<\infty,$$

for μ-a.e. $x\in X$. Let us summarize the preceding discussion in the following statement.

THEOREM 38. *Assume that the operator $M_{\mathbf{R}}$ is of weak type (p_0,p_0) for some $1<p_0<\infty$. Then, for any measure preserving transformations $S',T':X\to X$, the maximal operator $M'_{\mathbf{R}}$ associated to S', T' and $\mathbf{R}$ in the obvious way by*

$$M'_{\mathbf{R}}f(x)=\sup_{i\in\mathbb{N}}\frac{1}{\#R_i}\sum_{(k,l)\in R_i}f\bigl(S'^kT'^lx\bigr),$$

verifies $M'_{\mathbf{R}}f(x)<\infty$ for any $f\in L\log L_+(X,\mu)$.

REMARK 39. In case **Q** has infinite width, Theorem 30(ii) (see also Hagelstein and Stokolos [6]) shows the latter theorem is "the best we can expect".

Appendix A: Material from Orlicz spaces

For our purposes, we call an *Orlicz function* any function $\Phi : [0,\infty) \to [0,\infty)$ satisfying the following conditions:

(O1) Φ is convex and increasing;
(O2) $\Phi(0) = 0$;
(O3) the function $t \mapsto \Phi(t^{1/2})$ is concave.

It should immediately be noticed that any function verifying (O1)–(O3) as above also satisfies the following doubling condition:

(Δ_2) there exists $C > 0$ such that for any $0 \le t < \infty$, $\Phi(2t) \le C\Phi(t)$.

Given a Young function Φ and a measure space (X,μ), we define for every measurable $f : X \to \mathbb{R}$ a number

$$\|f\|_\Phi = \inf\left\{\alpha > 0 : \int_X \Phi\left(\frac{|f(x)|}{\alpha}\right) d\mu(x) \le 1\right\},$$

and we denote by $\mathscr{L}^\Phi(X)$ the collection of all measurable functions f on X for which $\|f\|_\Phi < \infty$, while $L^\Phi(X)$ denotes the space obtained by identifying, in $\mathscr{L}^\Phi(X)$, two almost everywhere equal functions. Here $f \in \mathscr{L}^\Phi(x)$ if and only if $\int_X \Phi(|f(x)|)\, d\mu(x) < \infty$. According to Rao and Ren [11, Proposition 3, p. 60; Theorem 10, p. 67], we find the following result:

THEOREM 40. *The space* $(L^\Phi(X), \|\cdot\|_\Phi)$ *is a Banach space.*

EXAMPLE 41. We denote by Φ_0 the Orlicz function defined on $[0,\infty)$ by

$$\Phi_0(t) = t(1 + \log_+ t),$$

and we let $L\log L(X) = L^{\Phi_0}(X)$ whenever (X,μ) is a measure space; we also let $\|f\|_{L\log L} = \|f\|_{\Phi_0}$ for $f \in L\log L(X)$. We readily notice that Φ_0 satisfies (Δ_2).

It will be convenient to use the following terminology, as in Stein [14, p. 154]: we will say that an operator T mapping $L^\Phi(X,\mu)$ into the space of $L^0(X,\mu)$ of measurable functions is

- *of (strong) type* (Φ,Φ) in case there exists a constant $C > 0$ such that for any $f \in L^\Phi(X,\mu)$ one has
$$\int_X \Phi\big(|Tf|\big)\, d\mu \le \int_X \Phi\big(C|f|\big)\, d\mu.$$
- *of weak type* (Φ,Φ) in case there exists a constant $C > 0$ such that for any $f \in L^\Phi(X,\mu)$ one has
$$\mu\big(\{x \in X : \big|T(x)\big| > \lambda\}\big) \le \int_X \Phi\left(\frac{C|f|}{\lambda}\right) d\mu.$$

Appendix B: Proof of the "transfer lemma"

In the sequel, we fix $n \geq 1$ an integer and we let $K_0 = [0,1)^n \subseteq \mathbb{R}^n$; consistently with this we also let $K_x = x + K_0$ for each $x \in \mathbb{R}^n$. Given $f \in L^1(\mathbb{R}^n)$ and $M > 0$, we let $D^M F \in L^1(\mathbb{R}^n)$ be defined by

$$D^M F(x) = M^n F(Mx).$$

We observe in particular that $\|D^M F\|_1 = \|F\|_1$.

A *regularizing* sequence is a sequence of smooth, compactly supported functions $(\rho_i) \subseteq L^1(\mathbb{R}^n)$ satisfying the following conditions:

(R1) $\rho_i(x) \geq 0$ for each $i \in \mathbb{N}$ and each $x \in \mathbb{R}^n$;
(R2) $\operatorname{supp} \rho_i \subseteq \{x \in \mathbb{R}^n : |x| \leq 2^{-i}\}$ for each $i \in \mathbb{N}$;
(R3) $\int_{\mathbb{R}^n} \rho_i = 1$ for each $i \in \mathbb{N}$.

Given a set $\Omega \subseteq \mathbb{R}^n$, we also let for $i \in \mathbb{N}$:

$$\Omega_i = \big\{x \in \Omega : \operatorname{dist}(x, \partial\Omega) > 2^{-i}\big\};$$

and we notice that if $f \in L^1(\mathbb{R}^n)$ is constant in Ω, then $\rho_i * f$ is constant in Ω_i.

Given $k \in \mathbb{Z}^n$ and $M \in \mathbb{N}^*$, we let $\mathbb{K}_k^M = \mathbb{Z}^n \cap MK_k$. We also denote by $\ell^1(\mathbb{Z}^n)$ the family of all summable functions on $\mathbb{Z}^n$. For each $k \in \mathbb{Z}^n$, we define a function $\delta_k \in \ell^1(\mathbb{Z}^n)$ by

$$\delta_k(l) = \begin{cases} 1 & \text{if } l = k, \\ 0 & \text{otherwise;} \end{cases}$$

while given $M \in \mathbb{N}^*$, we let $\alpha_M = M^{-n} \mathbb{1}_{\mathbb{K}_0^M}$. Given $\varphi \in \ell^1(\mathbb{Z}^n)$ and $M \in \mathbb{N}^*$, we define $D_M\varphi \in \ell^1(\mathbb{Z}^n)$ and $\Delta_M\varphi \in \ell^1(\mathbb{Z}^n)$ by

$$D_M\varphi = M^{-n} \sum_{k \in \mathbb{Z}^n} \varphi(k) \mathbb{1}_{\mathbb{K}_k^M} \quad \text{and} \quad \Delta_M\varphi = \sum_{k \in \mathbb{Z}^n} \varphi(k) \delta_{Mk}.$$

We easily observe that $\|\varphi\|_1 = \|D_M\varphi\|_1 = \|\Delta_M\varphi\|_1$.

The following result is our "transfer lemma".

LEMMA 42. *Let $\mathscr{F} \subseteq \ell^1(\mathbb{Z}^n)$ be a countable collection of summable functions on $\mathbb{Z}^n$ and define for each $\varphi \in \mathscr{F}$ a function $f_\varphi \in L^1(\mathbb{R}^n)$ by $f_\varphi = \sum_{k \in \mathbb{Z}^n} \varphi(k) \mathbb{1}_{K_k}$. The following assertions are equivalent:*

(A) *there exists $C = C(\mathscr{F}, n) > 0$ such that for any $g \in L^1(\mathbb{R}^n)$ and $\lambda > 0$, we have*

$$\Big|\Big\{\sup_{\varphi \in \mathscr{F}} f_\varphi * g > \lambda\Big\}\Big| \leq \frac{C}{\lambda} \|g\|_1;$$

(B) *there exists $C = C(\mathscr{F}, n) > 0$ such that for any $M \in \mathbb{N}^*$, $\psi \in \ell^1(\mathbb{Z}^n)$ and $\lambda > 0$ we have*

$$\#\Big\{\sup_{\varphi \in \mathscr{F}} D_M\varphi * \psi > \lambda\Big\} \leq \frac{C}{\lambda} \|\psi\|_1;$$

(C1) *there exists* $C = C(\mathscr{F}, n) > 0$ *such that for any* $\psi \in \ell^1(\mathbb{Z}^n)$ *and any* $\lambda > 0$, *we have*

$$\#\Big\{ \sup_{\varphi \in \mathscr{F}} \varphi * \psi > \lambda \Big\} \leq \frac{C}{\lambda} \|\psi\|_1;$$

(CM) *there exists* $C = C(\mathscr{F}, n) > 0$ *such that for any* $M \in \mathbb{N}^*$, $\psi \in \ell^1(\mathbb{Z}^n)$ *and* $\lambda > 0$, *we have*

$$\#\Big\{ \sup_{\varphi \in \mathscr{F}} \Delta_M \varphi * \psi > \lambda \Big\} \leq \frac{C}{\lambda} \|\psi\|_1.$$

We shall prove the following series of implications: (B) $\Rightarrow$ (A) $\Rightarrow$ (B) $\Rightarrow$ (C1) $\Rightarrow$ (CM) $\Rightarrow$ (B).

Proof that (B) *implies* (A). We will need the following observation. Given $k, l \in \mathbb{Z}^n$ we let $E_{k,l} := \{x \in \mathbb{R}^n : |K_k \cap (x - K_l)| > 0\}$ and $\mathbb{1}_{k,l} := \mathbb{1}_{E_{k,l}}$.

Claim 43. *For each* $k, l \in \mathbb{Z}^n$, *we have* $\mathbb{1}_{k,l} \leq \mathbb{1}_{K_{k+l}+K_0}$.

Proof. Assuming that $x \in \mathbb{R}^n$ satisfies $|K_k \cap (x - K_l)| > 0$, observe that we can find $y \in K_0$ for which $x - (l + y) \in K_k$. In particular we have $x \in K_{k+l} + y \subseteq K_{k+l} + K_0$. □

We now turn on to show that (B) implies (A). Begin by observing that the family

$$\mathscr{D} = \big\{ D^M f_\psi : M \in \mathbb{N}^*, \psi \in \ell^1\big(\mathbb{Z}^n\big) \big\}$$

is dense in $L^1(\mathbb{R}^n)$ (to see this, notice that the latter collection is that of all step functions in $L^1(\mathbb{R}^n)$ subordinate to a regular grid centered at 0 whose step is the inverse of a positive integer). It is then sufficient to show (B) for functions $g \in \mathscr{D}$.

On the other hand, taking $g = f_\psi$ with $\psi \in \ell^1(\mathbb{Z}^n)$, we compute for $\varphi \in \mathscr{F}$, $M \in \mathbb{N}^*$ and $x \in \mathbb{R}^n$:

$$\begin{aligned} f_\varphi * D^M f_\psi(x) &= \int_{\mathbb{R}^n} f_\varphi(y) f_\psi\big(M(x-y)\big) M^n \, dy \\ &= M^n \int_{\mathbb{R}^n} M^{-n} f_\varphi\big(M^{-1}(Mx - z)\big) f_\psi(z) \, dz \\ &= D^M\big(\big(D^{M^{-1}} f_\varphi\big) * f_\psi\big)(x); \end{aligned}$$

while on the other hand

$$D^{M^{-1}} f_\varphi(x) = M^{-n} \sum_{l \in \mathbb{Z}^n} \varphi(l) \mathbb{1}_{K_l}\big(M^{-1}x\big) = M^{-n} \sum_{l \in \mathbb{Z}^n} \varphi(l) \mathbb{1}_{MK_l}(x).$$

Yet we have

$$\begin{aligned} f_{D_M \varphi}(x) &= \sum_{k \in \mathbb{Z}^n} D_M \varphi(k) \mathbb{1}_{K_k}(x) \\ &= M^{-n} \sum_{k \in \mathbb{Z}^n} \sum_{l \in \mathbb{Z}^n} \varphi(l) \mathbb{1}_{MK_l}(k) \mathbb{1}_{K_k}(x) \end{aligned}$$

$$= M^{-n} \sum_{l\in\mathbb{Z}^n} \varphi(l) \sum_{k\in\mathbb{Z}^n} \mathbb{1}_{MK_l}(k)\mathbb{1}_{K_k}(x)$$
$$= M^{-n} \sum_{l\in\mathbb{Z}^n} \varphi(l)\mathbb{1}_{MK_l}(x),$$

so that finally

$$f_\varphi * D^M f_\psi = D^M(f_{D_M\varphi} * f_\psi).$$

Now fix $\eta > 0$ and $\psi \in \ell^1(\mathbb{Z}^n)$. Assuming that $x \in \mathbb{R}^n$ satisfies

$$f_{D_M\varphi} * f_\psi(x) > \eta$$

for some $\varphi \in \varphi$, observe that we have

$$\sum_{k,l\in\mathbb{Z}^n} \mathbb{1}_{K_{k,l}+K_0}(x) D_M\varphi(k)\psi(l) \geq \sum_{k,l\in\mathbb{Z}^n} \mathbb{1}_{k,l}(x) D_M\varphi(k)\psi(l) \tag{9}$$
$$= f_{D_M\varphi} * f_\psi(x) > \eta.$$

We now denote by $\mathscr{C}_n$ the collection of all subsets of $\{1, 2, \ldots, n\}$ and we observe that $\#\mathscr{C}_n = 2^n$. Given $x \in \mathbb{R}^n$ and $\gamma \in \mathscr{C}_n$, we let $[x]_\gamma$ satisfy $[x]_\gamma^{(i)} = [x^{(i)}] - 1$ in case $i \in \gamma$ and $[x]_\gamma^{(i)} = [x^{(i)}]$ otherwise. In particular, $[x]_\emptyset = [x]$. Observing that for each $\gamma \in \mathscr{C}_n$, we have

$$x \in K_{k+l} + K_0 \quad \text{for each } k, l \in \mathbb{Z}^n \text{ with } k + l = [x]_\gamma,$$

we write

$$\sum_{k,l\in\mathbb{Z}^n} \mathbb{1}_{K_{k,l}+K_0}(x) D_M\varphi(k)\psi(l) = \sum_{\gamma\in\mathscr{C}_n} \sum_{k\in\mathbb{Z}^n} D_M\varphi(k)\psi\big([x]_\gamma - k\big),$$

and infer from (9) that

$$D_M\varphi * \psi\big([x]_\gamma\big) = \sum_{k\in\mathbb{Z}^n} D_M\varphi(k)\psi\big([x]_\gamma - k\big) > \frac{\eta}{2^n},$$

for at least one $\gamma \in \mathscr{C}_n$.

Fix now $\lambda > 0$ and $\psi \in \ell^1(\mathbb{Z}^n)$. We define for each $M \in \mathbb{N}^*$:

$$B_M = \Big\{x \in \mathbb{R}^n : \sup_{\varphi\in\mathscr{F}} f_\varphi * D^M f_\psi(x) > \lambda\Big\}$$
$$= \Big\{x \in \mathbb{R}^n : \sup_{\varphi\in\mathscr{F}} D^M(f_{D_M\varphi} * f_\psi)(x) > \lambda\Big\}$$
$$= \Big\{x \in \mathbb{R}^n : \sup_{\varphi\in\mathscr{F}} f_{D_M\varphi} * f_\psi(Mx) > \frac{\lambda}{M^n}\Big\}$$
$$= M^{-1}\Big\{y \in \mathbb{R}^n : \sup_{\varphi\in\mathscr{F}} f_{D_M\varphi} * f_\psi(y) > \frac{\lambda}{M^n}\Big\}.$$

Calling B'_M the latter set, we now observe (applying what precedes to $\eta = \lambda/M^n$) that

$$B'_M \subseteq \bigcup_{\gamma\in\mathscr{C}_n} \left\{ x\in\mathbb{R}^n : \sup_{\varphi\in\mathscr{F}} D_M\varphi * \psi\big([x]_\gamma\big) > \frac{\lambda}{(2M)^n} \right\},$$

and we see by (B) that

$$\begin{aligned} \big|B'_M\big| &\le \sum_{\gamma\in\mathscr{C}_n} \#\left\{ k : \sup_{\varphi\in\mathscr{F}} D_M\varphi * \psi(k) > \frac{\lambda}{(2M)^n} \right\} \\ &\le M^n \frac{4^n C}{\lambda} \|\psi\|_1 = M^n \frac{C'}{\lambda} \|\psi\|_1. \end{aligned}$$

Hence,

$$|B_M| \le M^{-n} \big|B'_M\big| \le \frac{C'}{\lambda} \|f_\psi\|_1 = \frac{C'}{\lambda} \big\|D^M f_\psi\big\|_1.$$

This yields (A). □

Proof that (A) *implies* (B). Fix $M\in\mathbb{N}^*$, $\psi\in\ell^1(\mathbb{Z}^n)$ and compute for $\varphi\in\mathscr{F}$ and $j\in\mathbb{Z}^n$:

$$\begin{aligned} D_M\varphi * \psi(j) &= \sum_{l\in\mathbb{Z}^n} D_M\varphi(l)\psi(j-l) \\ &= M^{-n} \sum_{l\in\mathbb{Z}^n} \sum_{k\in\mathbb{Z}^n} \varphi(k) \mathbb{1}_{MK_k}(l) \psi(j-l) \\ &= M^{-n} \sum_{k\in\mathbb{Z}^n} \varphi(k) \sum_{l\in\mathbb{Z}^n} \psi(l) \mathbb{1}_{MK_k}(j-l) \\ &= M^{-n} \sum_{k\in\mathbb{Z}^n} \varphi(k) \sum_{l\in\mathbb{Z}^n} \psi(l) \mathbb{1}_{l+MK_k}(j). \end{aligned}$$

Define for each $i\in\mathbb{N}^*$ a function $g_i^M \in L^1(\mathbb{R}^n)$ by

$$g_i^M(x) = M^{-n} \sum_{k\in\mathbb{Z}^n} \psi(k) \rho_i\left(x - \frac{k}{M}\right).$$

We hence compute for $\varphi\in\mathscr{F}$ and $x\in\mathbb{R}^n$:

$$\begin{aligned} f_\varphi * g_i^M(x) &= \int_{\mathbb{R}^n} f_\varphi(y) g_i^M(x-y)\,dy \\ &= M^{-n} \sum_{k\in\mathbb{Z}^n} \varphi(k) \sum_{l\in\mathbb{Z}^n} \psi(l) \int_{\mathbb{R}^n} \mathbb{1}_{K_k}(y) \rho_i\left(x - y - \frac{l}{M}\right) dy \\ &= M^{-n} \sum_{k\in\mathbb{Z}^n} \varphi(k) \sum_{l\in\mathbb{Z}^n} \psi(l) \int_{\mathbb{R}^n} \rho_i(x-z) \mathbb{1}_{\frac{l}{M}+K_k}(z)\,dz \\ &= M^{-n} \sum_{k\in\mathbb{Z}^n} \varphi(k) \sum_{l\in\mathbb{Z}^n} \psi(l) \rho_i * \mathbb{1}_{\frac{l}{M}+K_k}(x). \end{aligned}$$

Assume now that $\lambda > 0$ is given and that $j \in \mathbb{Z}^n$ is such that for some $\varphi \in \mathscr{F}$ we have

$$D_M \varphi * \psi(j) = M^{-n} \sum_{k \in \mathbb{Z}^n} \varphi(k) \sum_{l \in \mathbb{Z}^n} \psi(l) \mathbb{1}_{l+MK_k}(j) > \lambda.$$

CLAIM 44. *If $k, l \in \mathbb{Z}^n$ are such that $\mathbb{1}_{l+MK_k}(j) = 1$, then for each $x \in (\frac{1}{M} K_j)_i$ we have*

$$\rho_i * \mathbb{1}_{\frac{l}{M}+K_k}(x) = 1.$$

Proof. Noticing to begin with that $\mathbb{1}_{l+MK_k}(j) = 1$ implies that $l \in j - MK_k$. Observe now that:

(1) $j - MK_k$ is a left-open cube,
(2) for each $x \in \frac{1}{M} K_j$, each coordinate of $Mx - j$ belongs to $[0, 1)$,
(3) l has integer coordinates;

we see that $l \in Mx - MK_k$ for each $x \in \frac{1}{M} K_j$.

On the other hand from $\mathbb{1}_{l+MK_k}(j) = 1$ we also infer that $j \in l + MK_k$; as $l + MK_k$ is a right-open cube in $\mathbb{R}^n$, we hence infer that $K_j \subseteq l + MK_k$, and thus

$$\frac{1}{M} K_j \subseteq \frac{l}{M} + K_k;$$

from which it follows that for any $x \in (\frac{1}{M} K_j)_i$ we have

$$\operatorname{dist}\left(x, \partial\left(\frac{l}{M} + K_k\right)\right) > 2^{-i}.$$

It then follows from property (R2) that $\rho_i * \mathbb{1}_{\frac{l}{M}+K_k}(x) = 1$. The claim is proved. □

According to the previous claim we thus have for each $x \in (\frac{1}{M} K_j)_i$:

$$\begin{aligned} f_\varphi * g_i^M(x) &= M^{-n} \sum_{k \in \mathbb{Z}^n} \varphi(k) \sum_{l \in \mathbb{Z}^n} \psi(l) \rho_i * \mathbb{1}_{\frac{l}{M}+K_k}(x) \\ &\geq M^{-n} \sum_{k \in \mathbb{Z}^n} \varphi(k) \sum_{l \in \mathbb{Z}^n} \psi(l) \mathbb{1}_{l+MK_k}(j) > \lambda. \end{aligned}$$

Letting $\varepsilon_\lambda = \{j \in \mathbb{Z}^n : \sup_{\varphi \in \mathscr{F}} \varphi * \psi(j) > \lambda\}$ we get from the preceding computations that

$$\bigcup_{j \in \varepsilon_\lambda} \left(\frac{1}{M} K_j\right)_i \subseteq \left\{x \in \mathbb{R}^n : \sup_{\varphi \in \mathscr{F}} \left(f_\varphi * g_i^M\right)(x) > \lambda\right\},$$

which yields in particular

$$\begin{aligned} M^{-n} \left(1 - \frac{1}{2^i M}\right)^n \# \varepsilon_\lambda &\leq \left|\left\{x \in \mathbb{R}^n : \sup_{\varphi \in \mathscr{F}} \left(f_\varphi * g_i^M\right)(x) > \lambda\right\}\right| \\ &\leq \frac{C}{\lambda} \left\|g_i^M\right\|_1 \leq M^{-n} \frac{C}{\lambda} \|\psi\|_1, \end{aligned}$$

and (B) follows since i is arbitrary. □

Proof that (B) *implies* (C1). This is obvious, for it it sufficient to take $M = 1$ in (B). □

Proof that (C1) *implies* (CM). We will make use of the following simple fact.

CLAIM 45. *Fix $M \in \mathbb{N}^*$ and $\psi_j \in \ell^1(\mathbb{Z}^n)$, $j \in \mathbb{K}_0^M$. Then for $k \in q + M\mathbb{Z}^n$, $q \in \mathbb{K}_0^M$, we have*

$$\sup_{\varphi\in\mathscr{F}} \sum_{j\in\mathbb{K}_0^M} \delta_j * (\Delta_M\varphi * \Delta_M\psi_j) = \sup_{\varphi\in\mathscr{F}} \Delta_M\varphi * \Delta_M\psi_q(k-q).$$

Proof. Fix $k \in \mathbb{Z}^n$ and choose $q \in \mathbb{K}_0^M$ and $r \in \mathbb{Z}^n$ for which $k = q + Mr$. For each $\varphi \in \mathscr{F}$ we have

$$\begin{aligned}
\sum_{j\in\mathbb{K}_0^M} \delta_j * (\Delta_M\varphi * \Delta_M\psi_j)(k) &= \sum_{j\in\mathbb{K}_0^M} \sum_{l\in\mathbb{Z}^n} \delta_j(l)\Delta_M\varphi * \Delta_M\psi_j(k-l) \\
&= \sum_{j\in\mathbb{K}_0^M} \Delta_M\varphi * \Delta_M\psi_j(k-j) \\
&= \sum_{\{j\in\mathbb{K}_0^M : k-j\in M\mathbb{Z}^n\}} \Delta_M\varphi * \Delta_M\psi_j(k-j) \\
&= \Delta_M\varphi * \Delta_M\psi_q(k-q),
\end{aligned}$$

in particular,

$$\sup_{\varphi\in\mathscr{F}} \sum_{j\in\mathbb{K}_0^M} \delta_j * (\Delta_M\varphi * \Delta_M\psi_j)(k) = \sup_{\varphi\in\mathscr{F}} (\Delta_M\varphi * \Delta_M\psi_q)(k-q);$$

the claim is proved. □

We now show that (C1) implies condition (CM); to that purpose, call $C > 0$ the constant appearing in (C1). It is then obvious that for each $M \in \mathbb{N}^*$ and each $\psi \in \ell^1(\mathbb{Z}^n)$ with $\operatorname{supp}\psi \subseteq M\mathbb{Z}^n$, we have for each $\lambda > 0$:

$$\#\Big\{ \sup_{\varphi\in\mathscr{F}} \Delta_M\varphi * \psi > \lambda \Big\} \le \frac{C}{\lambda}\|\psi\|_1.$$

Given $\psi \in \ell^1(\mathbb{Z}^n)$, $M \in \mathbb{N}^*$ and $j \in \mathbb{K}_0^M$, define $\psi_j^M \in \ell^1(\mathbb{Z}^n)$ by

$$\psi_j^M(k) = \psi(j + Mk).$$

Observe that $\sum_{j\in\mathbb{K}_0^M} \|\psi_j^M\|_1 = \|\psi\|_1$ and

$$\psi = \sum_{j\in\mathbb{K}_0^M} \delta_j * \Delta_M\psi_j.$$

In particular, we compute for $\lambda > 0$ and $\psi \in \ell^1(\mathbb{Z}^n)$, according to Claim 45:

$$\begin{aligned}
&\#\Big\{k \in \mathbb{Z}^n : \sup_{\varphi \in \mathscr{F}} \Delta_M \varphi * \psi(k) > \lambda\Big\} \\
&\quad = \sum_{q \in \mathbb{K}_0^M} \#\Big\{k \in q + M\mathbb{Z}^n : \sup_{\varphi \in \mathscr{F}} \Delta_M \varphi * \Delta_M \psi_q^M(k-q) > \lambda\Big\} \\
&\quad = \sum_{q \in \mathbb{K}_0^M} \#\Big\{l \in M\mathbb{Z}^n : \sup_{\varphi \in \mathscr{F}} \Delta_M \varphi * \Delta_M \psi_q^M(l) > \lambda\Big\} \\
&\quad \leq \frac{C}{\lambda} \sum_{q \in \mathbb{K}_0^M} \big\|\Delta_M \psi_q^M\big\|_1 \\
&\quad = \frac{C}{\lambda} \sum_{q \in \mathbb{K}_0^M} \big\|\psi_q^M\big\|_1 \\
&\quad = \frac{C}{\lambda} \|\psi\|_1;
\end{aligned}$$

which establishes (CM). □

Proof that (CM) *implies* (B). The following easy fact will be useful in the sequel.

Claim 46. *For any $\psi \in \ell^1(\mathbb{Z}^n)$, we have $\|\alpha_M * \psi\|_1 \leq \|\psi\|_1$.*

Proof. We easily compute

$$\|\alpha_M * \psi\|_1 = M^{-n} \sum_{k \in \mathbb{Z}^n} \Big| \sum_{l \in \mathbb{K}_0^M} \psi(k-l) \Big| \leq M^{-n} \sum_{l \in \mathbb{K}_0^M} \sum_{k \in \mathbb{Z}^n} \big|\psi(k-l)\big| = \|\psi\|_1.$$

The claim is proved. □

We will also need the following fact.

Claim 47. *For each $\phi \in \ell^1(\mathbb{Z}^n)$ and each $M \in \mathbb{N}^*$, we have $D_M \phi = \alpha_M * \Delta_M \phi$.*

Proof. To prove this claim, fix $j \in \mathbb{Z}^n$ and $k \in \mathbb{K}_j^M$. Compute then

$$\begin{aligned}
\alpha_M * \Delta_M \varphi(k) &= \sum_{l \in \mathbb{Z}^n} \alpha_M(l) \Delta_M \varphi(k-l), \\
&= M^{-n} \sum_{\{l \in \mathbb{K}_0^M : k-l \in M\mathbb{Z}^n\}} \varphi\Big(\frac{k-l}{M}\Big) \\
&= M^{-n} \varphi\Big(\frac{k-(k-jM)}{M}\Big) \\
&= M^{-n} \varphi(j) \\
&= D_M \varphi(k);
\end{aligned} \tag{10}$$

where (10) comes from the fact that the only $l \in \mathbb{K}_0^M$ for which $k - l \in M\mathbb{Z}^n$ is the vector $l = k - Mj$. □

We are now able to prove that condition (CM) implies condition (B): to that purpose assume that (CM) holds and compute for $\psi \in \ell^1(\mathbb{Z}^n)$ and $\lambda > 0$ (using (CM) together with the previous claims):

$$\#\Big\{\sup_{\varphi\in\mathscr{F}} D_M\varphi * \psi > \lambda\Big\} = \#\Big\{\sup_{\varphi\in\mathscr{F}} \Delta_M\varphi * (\alpha_M * \psi) > \lambda\Big\}$$
$$\leq \frac{C}{\lambda}\|\alpha_M * \psi\|_1 \leq \frac{C}{\lambda}\|\psi\|_1.$$

Hence, (B) is proved. □

Acknowledgments. Laurent Moonens would like to thank the Department of Mathematics of the University of Illinois at Urbana-Champaign, and especially Professor J. Rosenblatt, for their warm hospitality during the academic year 2009–2010.

The authors would like the referee for a careful reading of the paper, and the numerous suggestions which improved the quality of the present text.

References

[1] P. Avramidou, *On certain weighted moving averages and their differentiation analogues*, New York J. Math. **12** (2006), 19–37. MR 2217161

[2] A. Bellow, R. Jones and J. Rosenblatt, *Convergence for moving averages*, Ergodic Theory Dynam. Systems **10** (1990), no. 1, 43–62. MR 1053798

[3] A. Bellow and R. L. Jones, *A Banach principle for L^∞*, Adv. Math. **120** (1996), no. 1, 155–172. MR 1392277

[4] A.-P. Calderón, *Ergodic theory and translation-invariant operators*, Proc. Natl. Acad. Sci. USA **59** (1968), 349–353. MR 0227354

[5] R. P. Dilworth, *A decomposition theorem for partially ordered sets*, Ann. of Math. (2) **51** (1950), 161–166. MR 0032578

[6] P. Hagelstein and A. Stokolos, *Weak type inequalities for ergodic strong maximal operators*, Acta Sci. Math. (Szeged) **76** (2010), 365–379. MR 2789679

[7] P. Hagelstein and A. Stokolos, *Weak type inequalities for maximal operators associated to double ergodic sums*, New York J. Math. **17** (2011), 233–250. MR 2781915

[8] B. Jessen, J. Marcinkiewicz and A. Zygmund, *Note on the differentiability of multiple integrals*, Fund. Math. **25** (1935), 217–234.

[9] A. Nagel and E. M. Stein, *On certain maximal functions and approach regions*, Adv. in Math. **54** (1984), no. 1, 83–106. MR 0761764

[10] D. S. Ornstein and B. Weiss, *Ergodic theory of amenable group actions. I. The Rohlin lemma*, Bull. Amer. Math. Soc. (N.S.) **2** (1980), no. 1, 161–164. MR 0551753

[11] M. M. Rao and Z. D. Ren, *Theory of Orlicz spaces*, Monographs and Textbooks in Pure and Applied Mathematics, vol. 146, Marcel Dekker Inc., New York, 1991. MR 1113700

[12] J. M. Rosenblatt and M. Wierdl, *A new maximal inequality and its applications*, Ergodic Theory Dynam. Systems **12** (1992), no. 3, 509–558. MR 1182661

[13] S. Sawyer, *Maximal inequalities of weak type*, Ann. of Math. (2) **84** (1966), 157–174. MR 0209867

[14] E. M. Stein, *On limits of seqences of operators*, Ann. of Math. (2) **74** (1961), 140–170. MR 0125392

[15] A. M. Stokolos, *On the differentiation of integrals of functions from* $L\varphi(L)$, Studia Math. **88** (1988), no. 2, 103–120. MR 0931036
[16] A. Zygmund, *An individual ergodic theorem for non-commutative transformations*, Acta Sci. Math. Szeged **14** (1951), 103–110. MR 0045948
[17] A. Zygmund, *Trigonometric series*, 2nd ed. vols. I, II, Cambridge Univ. Press, New York, 1959. MR 0107776

Laurent Moonens, Université Paris-Sud, Laboratoire de Mathématiques, Bâtiment 425, 91405 Orsay Cedex, France

E-mail address: Laurent.Moonens@math.u-psud.fr

Joseph M. Rosenblatt, Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

E-mail address: rosnbltt@illinois.edu