ISRAEL JOURNAL OF MATHEMATICS 135 (2003), 1–27

OSCILLATION IN ERGODIC THEORY: HIGHER DIMENSIONAL RESULTS

BY

ROGER L. JONES*

Department of Mathematics, DePaul University
2219 N. Kenmore, Chicago, IL 60614, USA
e-mail: rjones@condor.depaul.edu

AND

JOSEPH M. ROSENBLATT†

Department of Mathematics, University of Illinois at Urbana
Urbana, IL 61801, USA
e-mail: jrsnbltt@math.uiuc.edu

AND

MÁTÉ WIERDL‡

Department of Mathematical Sciences, University of Memphis
Memphis, TN 38152, USA
e-mail: mw@csi.hu

ABSTRACT

In this paper we continue our investigations of square function inequalities. The results in [9] are primarily one dimensional, and here we extend all the results to multi-dimensional averages. Our basic tool is still a comparison of the ergodic averages with various dyadic (reversed) martingales, but the Fourier transform arguments are replaced by more geometric almost orthogonality arguments.

The results imply the pointwise ergodic theorem for the action of commuting measure preserving transformations, and give additional information such as control of the number of upcrossings of the ergodic averages. Related differentiation results are also discussed.

* R. Jones is partially supported by NSF Grant DMS-9531526.
† J. Rosenblatt is partially supported by NSF Grant DMS-9705228.
‡ M. Wierdl is partially supported by NSF Grant DMS-9801602.
Received August 17, 2000

1. Introduction, main results for cubes

Let (X, Σ, m, τ) denote a dynamical system with (X, Σ, m) a σ finite measure space and $\tau = \tau_y$ a measurable, measure preserving $\mathbb{Z}^d$ action on X.

For a region $A \subset \mathbb{Z}^d$ and $f : X \to \mathbb{R}$, $M_A f(x)$ denotes the average operator over A:

$$M_A f(x) = \frac{1}{|A|} \sum_{y \in A} f(\tau_y x).$$

The ergodic theorem of Wiener says that if (A_n) is a sequence of cubes in $\mathbb{Z}^d$ with $|A_n| \to \infty$ and each A_n containing the origin, then the averages $M_{A_n} f$ converge almost everywhere for $f \in L^1$, and the maximal function $\sup_n |M_{A_n} f|$ is weak type $(1,1)$ and type (p,p) for $p > 1$.

In the particular case, when $X = \mathbb{Z}^d$, m is the counting measure, and $\tau_y x = x + y$, the maximal function $\sup_n |M_{A_n} f|$ is just the Hardy–Littlewood maximal function. By the transference principle of Calderón, the Hardy–Littlewood maximal inequalities imply those for the ergodic maximal function. This remark applies to all the inequalities to be proved in this paper: the real work is to be done on the acting group, and a routine argument transfers the result to any dynamical system.

An important observation, that will be very useful in obtaining the results to follow, is that there is a square function that connects a (reverse) martingale with averages over cubes. Let $X = \mathbb{Z}^d$. For $n = 0, 1, \ldots$, let σ_n be the n-th dyadic σ algebra; that is, σ_n is generated by the dyadic cubes,

$$[s_1 2^n, (s_1 + 1)2^n) \times [s_2 2^n, (s_2 + 1)2^n) \times \cdots \times [s_d 2^n, (s_d + 1)2^n)$$

where the s_i's are integers. Denoting by $\mathcal{E}_n$ the expectation with respect to σ_n, we have

THEOREM A: *For each n, let A_n be a cube in $\mathbb{Z}^d$ with side-length 2^n and containing the origin. Then*

$$\left\| \left(\sum_n |M_{A_n} f - \mathcal{E}_n f|^2 \right)^{1/2} \right\|_{\ell^p} \leq c_p \cdot \|f\|_{\ell^p} \quad (1 < p \leq 2),$$

and

$$\left| \left\{ \left(\sum_n |M_{A_n} f - \mathcal{E}_n f|^2 \right)^{1/2} > \alpha \right\} \right| \leq \frac{c}{\alpha} \cdot \|f\|_{\ell^1}.$$

Remark: The constants c, c_p in Theorem A as well as in the other results of this paper may depend on the dimension d. Also, the value of the constants c, C are not fixed, and it may change even in the same argument.

Theorem A allows us to use results known for reverse martingales to obtain results for averages over cubes. In fact, we will prove a stronger form of Theorem A. This strengthened version allows the sequence A_n to depend on the point x and on the function f. Denote

$$H_n = [-2^{n+1}, 2^{n+1})^d$$

and let $\mathcal{A}_n$ contain all cubes with side-length between 2^{n-1} and 2^n which are contained in H_n. Note that for each point x, the atom H of σ_n containing the point x satisfies $H - x \in \mathcal{A}_n$; in other words, for some $A \in \mathcal{A}_n$ we have $\mathcal{E}_n f(x) = M_A f(x)$.

THEOREM A′: *We have*

$$\left\| \left(\sum_n \sup_{A \in \mathcal{A}_n} |M_A f - \mathcal{E}_n f|^2 \right)^{1/2} \right\|_{\ell^p} \leq c_p \cdot \|f\|_{\ell^p} \quad (1 < p \leq 2),$$

and

$$\left| \left\{ \left(\sum_n \sup_{A \in \mathcal{A}_n} |M_A f - \mathcal{E}_n f|^2 \right)^{1/2} > \alpha \right\} \right| \leq \frac{c}{\alpha} \cdot \|f\|_{\ell^1}.$$

An immediate corollary is

COROLLARY: *We have*

$$\left\| \left(\sum_n \sup_{A, B \in \mathcal{A}_n} |M_A f - M_B f|^2 \right)^{1/2} \right\|_{\ell^p} \leq c_p \cdot \|f\|_{\ell^p} \quad (1 < p \leq 2),$$

and

$$\left| \left\{ \left(\sum_n \sup_{A, B \in \mathcal{A}_n} |M_A f - M_B f|^2 \right)^{1/2} > \alpha \right\} \right| \leq \frac{c}{\alpha} \cdot \|f\|_{\ell^1}.$$

Let us list a few interesting consequences of Theorem A′. In Examples 1.1–1.5, for each $x \in \mathbb{Z}^d$ and function f, (A_n) is a sequence of cubes with $A_n \in \mathcal{A}_n$ for each n. In other words, the side-length of the cubes forms a lacunary sequence. Because of this restriction on the sequence (A_n), some of the results quoted in the examples are not in their most general form. For example, the Hardy–Littlewood maximal inequalities are basically derived only for averages over dyadic cubes. In case of the Hardy–Littlewood maximal inequalities, this more restrictive result simply implies the general form of the maximal inequalities, but in the other cases we will see how Theorem B below can be used to obtain the most general results for averages over cubes.

Note also that all the inequalities mentioned for the averages M_{A_n} are valid in any dynamical system by the transference principle.

Example 1.1 (Hardy–Littlewood maximal inequalities): We can assume that the function f is nonnegative. Then the Hardy–Littlewood inequalities follow from the pointwise estimate

$$\begin{aligned}\sup_n M_{A_n} f &\leq \sup_n \mathcal{E}_n f + \sup_n |M_{A_n} f - \mathcal{E}_n f| \\ &\leq \sup_n \mathcal{E}_n f + \Big(\sum_n |M_{A_n} f - \mathcal{E}_n f|^2\Big)^{1/2}\end{aligned}$$

and from the maximal inequalities for reverse martingales.

Example 1.2 (Square-function inequalities): We have the inequalities

$$\left\| \Big(\sum_n |M_{A_n} f - M_{A_{n+1}} f|^2\Big)^{1/2} \right\|_{\ell^p} \leq c_p \cdot \|f\|_{\ell^p} \quad (1 < p \leq 2),$$

and

$$\left| \left\{ \Big(\sum_n |M_{A_n} f - M_{A_{n+1}} f|^2\Big)^{1/2} > \alpha \right\} \right| \leq \frac{c}{\alpha} \cdot \|f\|_{\ell^1}.$$

This follows from the pointwise estimate

$$|M_{A_n} f - M_{A_{n+1}} f| \leq |\mathcal{E}_n f - \mathcal{E}_{n+1} f| + |M_{A_n} f - \mathcal{E}_n f| + |M_{A_{n+1}} f - \mathcal{E}_{n+1} f|$$

and from the inequalities for $(\sum_n |\mathcal{E}_n f - \mathcal{E}_{n+1} f|^2)^{1/2}$.

Example 1.3 (Oscillation inequalities): Let $n_1 < n_2 < \cdots$ be a sequence of integers. We have the inequalities

$$\left\| \Big(\sum_k \sup_{n_k \leq n \leq m < n_{k+1}} |M_{A_n} f - M_{A_m} f|^2\Big)^{1/2} \right\|_{\ell^p} \leq c_p \cdot \|f\|_{\ell^p} \quad (1 < p \leq 2),$$

and

$$\left| \left\{ \Big(\sum_k \sup_{n_k \leq n \leq m < n_{k+1}} |M_{A_n} f - M_{A_m} f|^2\Big)^{1/2} > \alpha \right\} \right| \leq \frac{c}{\alpha} \cdot \|f\|_{\ell^1}.$$

Note that the above inequalities in a dynamical system directly imply almost everywhere convergence of the averages $M_{A_n} f$. The inequalities follow from the pointwise estimate

$$|M_{A_n} f - M_{A_m} f| \leq |\mathcal{E}_n f - \mathcal{E}_m f| + |M_{A_n} f - \mathcal{E}_n f| + |M_{A_m} f - \mathcal{E}_m f|$$

and from the well-known inequalities for $(\sum_k \sup_{n_k \leq n \leq m < n_{k+1}} |\mathcal{E}_n f - \mathcal{E}_m f|^2)^{1/2}$ (cf. [9, Theorem 6.1]).

Example 1.4 (Jump inequalities): First, recall the concept of 'λ-jumps'. For $\lambda > 0$, we define the number of λ-jumps, $J((M_{A_n})), \lambda) = J((M_{A_n}(f,x)), \lambda)$, as the largest N for which there are indices $n_1 < n_2 < \cdots < n_N$ with

$$|M_{A_{n_i}} f(x) - M_{A_{n_{i+1}}} f(x)| > \lambda, \quad i = 1, 2, \ldots, N-1.$$

We have the inequalities

$$\|\lambda \cdot (J((M_{A_n}), \lambda))^{1/2}\|_{\ell^p} \leq c_p \cdot \|f\|_{\ell^p} \quad (1 < p \leq 2),$$

and

$$|\{\lambda \cdot (J((M_{A_n}), \lambda))^{1/2} > \alpha\}| \leq \frac{c}{\alpha} \cdot \|f\|_{\ell^1}.$$

Taking $\alpha = \lambda \cdot N^{1/2}$ in the weak $(1,1)$ inequality, we obtain

$$|\{J((M_{A_n}), \lambda) > N\}| \leq \frac{c}{\lambda N^{1/2}} \cdot \|f\|_{\ell^1}.$$

To obtain the inequalities, estimate first as

$$\lambda \sqrt{J((M_{A_n}), \lambda)} \leq \lambda \sqrt{J((M_{A_n} - \mathcal{E}_n), \lambda/2)} + \lambda \sqrt{J((\mathcal{E}_n), \lambda/2)}.$$

The first term above is handled with the estimate

$$\lambda \sqrt{J((M_{A_n} - \mathcal{E}_n), \lambda/2)} \leq 2 \cdot \sqrt{\sum_n |M_{A_n}(f,x) - \mathcal{E}_n f(x)|^2};$$

The second term is handled by the appropriate inequalities for reverse martingales (cf. [9, Section 6]).

Example 1.5 (Variation inequalities): For any $\varrho > 2$ we have the inequalities

$$\left\| \left(\sup_{(n_i)} \sum_i |M_{A_{n_i}} f - M_{A_{n_{i+1}}} f|^\varrho \right)^{1/\varrho} \right\|_{\ell^p} \leq c_{p,\varrho} \cdot \|f\|_{\ell^p} \quad (1 < p \leq 2),$$

and

$$\left| \left\{ \left(\sup_{(n_i)} \sum_i |M_{A_{n_i}} f - M_{A_{n_{i+1}}} f|^\varrho \right)^{1/\varrho} > \alpha \right\} \right| \leq \frac{c_\varrho}{\alpha} \cdot \|f\|_{\ell^1}.$$

Above, the supremum is taken over all increasing sequences $n_1 < n_2 < \cdots$.

The inequalities follow from the pointwise estimate

$$\begin{aligned}
\Big(\sum_i |M_{A_{n_i}}f - M_{A_{n_{i+1}}}f|^\varrho\Big)^{1/\varrho}
&\le \Big(\sum_i |(M_{A_{n_i}}f - \mathcal{E}_{n_i}f) + (\mathcal{E}_{n_{i+1}}f - M_{A_{n_{i+1}}}f)|^\varrho\Big)^{1/\varrho} \\
&\quad + \Big(\sum_i |\mathcal{E}_{n_i}f - \mathcal{E}_{n_{i+1}}f|^\varrho\Big)^{1/\varrho} \\
&\le \Big(\sum_i |(M_{A_{n_i}}f - \mathcal{E}_{n_i}f) + (\mathcal{E}_{n_{i+1}}f - M_{A_{n_{i+1}}}f)|^2\Big)^{1/2} \\
&\quad + \Big(\sum_i |\mathcal{E}_{n_i}f - \mathcal{E}_{n_{i+1}}f|^\varrho\Big)^{1/\varrho} \\
&\le 2^{1/2}\cdot\Big(\sum_n (M_{A_n}f - \mathcal{E}_nf)^2\Big)^{1/2} \\
&\quad + \Big(\sum_i |\mathcal{E}_{n_i}f - \mathcal{E}_{n_{i+1}}f|^\varrho\Big)^{1/\varrho}
\end{aligned}$$

and from the inequalities for $(\sup_{(n_i)}\sum_i |\mathcal{E}_{n_i}f - \mathcal{E}_{n_{i+1}}f|^\varrho)^{1/\varrho}$ (cf. [9, Theorems 6.2 and 6.4]).

So far, the sequence (A_n) of cubes has been restricted in the sense that $A_n \in \mathcal{A}_n$, that is, roughly speaking, the sequence of side-lengths forms a lacunary sequence. Of course, one would like to consider more general sequences (A_t) of cubes, where for a given n there might be *several* indices t with $A_t \in \mathcal{A}_n$. A natural example would be a sequence (A_t) of cubes with A_t of side-length t and containing the origin. In this case, $A_t \in \mathcal{A}_n$ for $2^{n-1} \le t < 2^n$. While the maximal function $\sup_t |M_{A_t}f|$ is bounded for such a sequence of cubes, the oscillatory operators we consider in this paper may fail to be bounded even in dimension $d = 1$ without further assumptions on (A_t).

To illustrate the problem one may encounter, consider the sequence (A_t) of intervals defined by

$$A_t = \begin{cases} [0, 2^n) & \text{if } t \text{ even}, \\ (-2^n, 0] & \text{if } t \text{ odd}, \end{cases} \qquad 2^n \le t < 2^{n+1}$$

and define the function $f_n : \mathbb{Z} \to \mathbb{R}$ as

$$f_n(x) = \begin{cases} 1 & \text{if } 0 \le x < 2^n, \\ -1 & \text{if } -2^n \le x < 0. \end{cases}$$

Clearly, we have $\|f_n\|_{\ell^2} = 2^{(n+1)/2}$. On the other hand, we have

$$\Big\|\Big(\sum_t (M_{A_t}f_n - M_{A_{t+1}}f_n)^2\Big)^{1/2}\Big\|_{\ell^2} > 2^{n-1},$$

which means that the square-function $(\sum_t (M_{A_t}f_n - M_{A_{t+1}}f_n)^2)^{1/2}$ is not a bounded $\ell^2 \to \ell^2$ operator.

To see the above inequality, note that for $2^n \le t < 2^{n+1}$ and $x \in [-2^n/4, 2^n/4)$ we have $|M_{A_t}f_n(x) - M_{A_{t+1}}f_n(x)| > 1$. As a consequence, we have

$$\begin{aligned}\int_{\mathbb{Z}} \sum_t (M_{A_t}f_n - M_{A_{t+1}}f_n)^2 &= \sum_t \int_{\mathbb{Z}} (M_{A_t}f_n - M_{A_{t+1}}f_n)^2\\ &\ge \sum_{2^n \le t < 2^{n+1}} \int_{\mathbb{Z}} (M_{A_t}f_n - M_{A_{t+1}}f_n)^2\\ &> 2^n \cdot 2^n/2,\end{aligned}$$

proving the inequality.

The further assumption we make on the sequence (A_t) is that if A_t and A_s are both in $\mathcal{A}_n$ for some n, then one of them contains the other. In other words, for every n the sets A_t with $A_t \in \mathcal{A}_n$ are *nested.* We'd like to emphasize the distinction between the set $\mathcal{A}_n$ and the sets A_t with $A_t \in \mathcal{A}_n$. Only the latter one, for each x, forms a nested sequence.

The conditions we use are summarized in the following definition:

Definition 1.6: We call the sequence (A_t) of cubes **regular** iff they satisfy the following conditions:

- for every n and $t \in [2^{n-1}, 2^n)$, we have $A_t \in \mathcal{A}_n$;
- $|A_s| \le |A_t|$ if $s \le t$;
- if $2^{n-1} \le s \le t < 2^n$ for some n, then $A_s \subset A_t$.

The next theorem enables us to extend all the results for 'lacunary' A_n's to an arbitrary regular sequence of A_t's.

THEOREM B: *Suppose that for each x and f, the sequence of cubes (A_t) is regular. Then we have the inequalities*

$$\Big\|\sup_{(t_i)}\Big(\sum_n \sum_{\substack{i\\ A_{t_i}, A_{t_{i+1}} \in \mathcal{A}_n}} |M_{A_{t_i}}f - M_{A_{t_{i+1}}}f|^2\Big)^{1/2}\Big\|_{\ell^p} \le c_p \cdot \|f\|_{\ell^p} \quad (1 < p \le 2),$$

and

$$\Big|\Big\{\Big(\sum_n \sup_{(t_i)} \sum_{\substack{i\\ A_{t_i}, A_{t_{i+1}} \in \mathcal{A}_n}} |M_{A_{t_i}}f - M_{A_{t_{i+1}}}f|^2\Big)^{1/2} > \alpha\Big\}\Big| \le \frac{c}{\alpha} \cdot \|f\|_{\ell^1}.$$

In the above, the supremum is taken over all increasing sequences $t_1 < t_2 < \cdots$.

All the applications of Theorem B are based on the following observation: Suppose (A_t) is a regular sequence of cubes and consider, say, the operator

$$\Big(\sum_i |M_{A_{t_i}} f - M_{A_{t_{i+1}}} f|^2\Big)^{1/2}.$$

For each index i, either both A_{t_i} and $A_{t_{i+1}}$ are from the same set $\mathcal{A}_n$, or $A_{t_i} \in \mathcal{A}_n$ and $A_{t_{i+1}} \in \mathcal{A}_m$ for some $n < m$. Denoting

$$\mathcal{S} = \{i \mid A_{t_i}, A_{t_{i+1}} \in \mathcal{A}_n \text{ for some } n\},$$
$$\mathcal{L} = \{i \mid A_{t_i} \in \mathcal{A}_n,\ A_{t_{i+1}} \in \mathcal{A}_m \text{ for some } n < m\},$$

we estimate as

$$\Big(\sum_i |M_{A_{t_i}} f - M_{A_{t_{i+1}}} f|^2\Big)^{1/2} \le \Big(\sum_{i\in\mathcal{S}} |M_{A_{t_i}} f - M_{A_{t_{i+1}}} f|^2\Big)^{1/2} + \Big(\sum_{i\in\mathcal{L}} |M_{A_{t_i}} f - M_{A_{t_{i+1}}} f|^2\Big)^{1/2}$$

Typically, we refer to the version of a given operator we get by restricting the indices i to $\mathcal{S}$ as the "short" version, and restricting the indices to $\mathcal{L}$ we obtain the "long" version of the operator. Hence the first term in the above sum is the short version of $(\sum_i |M_{A_{t_i}} f - M_{A_{t_{i+1}}} f|^2)^{1/2}$, and the second term is the long version.

We have already shown how the long versions of an operator can be handled using Theorem A$'$, and the short versions are handled by simply invoking Theorem B.

In the following theorems, we assume that for each point x and function f, the sequence (A_t) is regular. We omit the proofs since, with only minor modifications, they are very much like the earlier proofs. The basic idea, like before, is to consider the short and long versions of the operator separately.

THEOREM 1.7: *Let $v_1 < v_2 < \cdots$ be a sequence of positive numbers, and for each x and f let t_k be such that $|A_{t_k}| = v_k$. Then we have the inequalities*

$$\Big\|\Big(\sum_k |M_{A_{t_k}} f - M_{A_{t_{k+1}}} f|^2\Big)^{1/2}\Big\|_{\ell^p} \le c_p \cdot \|f\|_{\ell^p} \quad (1 < p \le 2),$$

and

$$\Big|\Big\{\Big(\sum_k |M_{A_{t_k}} f - M_{A_{t_{k+1}}} f|^2\Big)^{1/2} > \alpha\Big\}\Big| \le \frac{c}{\alpha} \cdot \|f\|_{\ell^1}.$$

The constants do not depend on the sequence (v_k).

THEOREM 1.8: *Let* $v_1 < v_2 < \cdots$ *be a sequence of positive numbers. Then we have the inequalities*

$$\left\| \left(\sum_k \sup_{v_k \le |A_s| \le |A_t| < v_{k+1}} |M_{A_s} f - M_{A_t} f|^2 \right)^{1/2} \right\|_{\ell^p} \le c_p \cdot \|f\|_{\ell^p} \quad (1 < p \le 2),$$

and

$$\left| \left\{ \left(\sum_k \sup_{v_k \le |A_s| \le |A_t| < v_{k+1}} |M_{A_s} f - M_{A_t} f|^2 \right)^{1/2} > \alpha \right\} \right| \le \frac{c}{\alpha} \cdot \|f\|_{\ell^1}.$$

The constants do not depend on the sequence (v_k).

THEOREM 1.9:

$$\left\| \lambda \cdot \left(J((M_{A_t}), \lambda) \right)^{1/2} \right\|_{\ell^p} \le c_p \cdot \|f\|_{\ell^p} \quad (1 < p \le 2),$$

and

$$|\{\lambda \cdot (J((M_{A_t}), \lambda))^{1/2} > \alpha\}| \le \frac{c}{\alpha} \cdot \|f\|_{\ell^1}.$$

COROLLARY:

$$|\{J((M_{A_t}), \lambda) > N\}| \le \frac{c}{\lambda N^{1/2}} \cdot \|f\|_{\ell^1}.$$

THEOREM 1.10: *For every* $\varrho > 2$, *we have*

$$\left\| \left(\sup_{(t_i)} \sum_i |M_{A_{t_i}} f - M_{A_{t_{i+1}}} f|^\varrho \right)^{1/\varrho} \right\|_{\ell^p} \le c_{p,\varrho} \cdot \|f\|_{\ell^p} \quad (1 < p \le 2),$$

and

$$\left| \left\{ \left(\sup_{(t_i)} \sum_i |M_{A_{t_i}} f - M_{A_{t_{i+1}}} f|^\varrho \right)^{1/\varrho} > \alpha \right\} \right| \le \frac{c_\varrho}{\alpha} \cdot \|f\|_{\ell^1}.$$

Above, the supremum is taken over all increasing sequences $t_1 < t_2 < \cdots$.

In this introductory section we considered averages over cubes only. In Section 5, we explain that our methods can be used to obtain analogs of Theorems 1.7–1.10 for other regions such as balls or even certain nonisotropic dilates of a fixed region. In Section 6, we discuss the differentiation analogs of our theorems in $\mathbb{R}^d$.

2. An almost orthogonality lemma

The proofs of Theorems A′ and B consist of proving the $(2,2)$ and the weak $(1,1)$ inequalities; the (p,p) inequalities follow by interpolation.

The proofs of both the $(2,2)$ and the weak $(1,1)$ inequalities rely on the following simple almost orthogonality result.

LEMMA 2.1: *Suppose $(\mathbb{S}_n)_{n\in\mathbb{Z}}$ is a sequence of subadditive ($|\mathbb{S}(f(x)+g(x))| \le |\mathbb{S}f(x)| + |\mathbb{S}g(x)|$) operators on L^2 in some σ-finite measure space. Let $(u_n)_{n\in\mathbb{Z}}$ and $(v_n)_{n\in\mathbb{Z}}$ be two sequences of L^2 functions. We assume that there is a sequence $(\sigma(j))_{j\in\mathbb{Z}}$ of positive numbers with $w = \sum_j \sigma(j) < \infty$ such that*

$$\|\mathbb{S}_k u_n\|_2 \le \sigma(n-k)\|v_n\|_2$$

for every n, k.

Then we have

$$\sum_k \left\| \sup_{j,m} |\mathbb{S}_k \sum_{j\le n\le m} u_n| \right\|_2^2 \le w^2 \cdot \sum_n \|v_n\|_2^2,$$

where the supremum $\sup_{j,m}$ is taken over all j,m with $j \le m$.

Furthermore, if the $\mathbb{S}$ are continuous, $f = \sum_n u_n$, $\sum_n \|v_n\|_2^2 \le C\|f\|_2^2$, then

$$\sum_k \|\mathbb{S}_k f\|_2^2 \le C \cdot w^2 \cdot \|f\|_2^2.$$

Proof: First note, using the subadditivity of the $\mathbb{S}$, that

$$\begin{aligned}\sup_{j,m} |\mathbb{S}_k \sum_{j\le n\le m} u_n| &\le \sup_{j,m} \sum_{j\le n\le m} |\mathbb{S}_k u_n| \\ &= \sum_n |\mathbb{S}_k u_n|,\end{aligned}$$

hence, by the triangle inequality for the ℓ^2 norm,

$$\left\| \sup_{j,m} |\mathbb{S}_k \sum_{j\le n\le m} u_n| \right\|_2 \le \sum_n \|\mathbb{S}_k u_n\|_2.$$

Because of the assumption $\|\mathbb{S}_k u_n\|_2 \le \sigma(n-k)\|v_n\|_2$, setting $b_n = \|v_n\|_2$, it is enough to prove

$$\sum_k \left(\sum_n \sigma(n-k) b_n\right)^2 \le \left(\sum_n \sigma(n)\right)^2 \cdot \sum_n b_n^2$$

But this is equivalent to the well known inequality for the ℓ^2 norm of the convolution of two sequences, $\|\sigma * b\|_{\ell^2} \le \|\sigma\|_{\ell^1} \cdot \|b\|_{\ell^2}$. ∎

3. Proof of the L^2 inequalities

Proof of Theorem A': Let us write $f = \sum_n d_n$ where $d_1 = f - \mathcal{E}_1 f$ and $d_n = \mathcal{E}_{n-1} f - \mathcal{E}_n f$ for $n > 1$. Then we have $\|f\|_{\ell^2}^2 = \sum_n \|d_n\|_{\ell^2}^2$.

Let A_n be such that $\sup_{A \in \mathcal{A}_n} |M_A f(x) - \mathcal{E}_n f(x)| = |M_{A_n} f(x) - \mathcal{E}_n f(x)|$.

Taking $\mathbb{S}_k$ defined by $\mathbb{S}_k g(x) = M_{A_k} g(x) - \mathcal{E}_k g(x)$, $u_n = v_n = d_n$ and $\sigma(j) = c \cdot 2^{-|j|/2}$ in the almost orthogonality lemma, Lemma 2.1, we see it is enough to prove

$$\|(M_{A_k} - \mathcal{E}_k) d_n\|_{\ell^2}^2 \leq c \cdot 2^{-|n-k|} \|d_n\|_{\ell^2}^2. \tag{3.1}$$

In order to prove (3.1), let us first assume $n > k$. Since in this case $\mathcal{E}_k d_n = d_n$, it is enough to prove

$$\|M_{A_k} d_n - d_n\|_2^2 \leq c \cdot 2^{k-n} \|d_n\|_2^2. \tag{3.2}$$

Let us denote by $\mathbb{D}_n$ the set of all atoms of the nth dyadic σ algebra σ_n, and write

$$\begin{aligned} \|M_{A_k} d_n - d_n\|_2^2 &= \int_{\mathbb{Z}^d} |M_{A_k} d_n - d_n|^2 \\ &= \sum_{H \in \mathbb{D}_{n-1}} \int_H |M_{A_k} d_n - d_n|^2. \end{aligned}$$

We want to estimate $\int_H |M_{A_k} d_n - d_n|^2$ for $H \in \mathbb{D}_{n-1}$. Since d_n is constant on the smaller atom H, we have $(M_{A_k} d_n - d_n)(x) = 0$ if $A_k + x \subset H$. It follows that $(M_{A_k} d_n - d_n)(x)$ may be nonzero only if $A_k + x$ intersects the boundary ∂H of H. Since $A_k \subset H_k$, the measure of those x in H for which $A_k + x \cap \partial H \neq \emptyset$ is at most a constant multiple of $2^{(d-1)n} \cdot 2^k$. Denoting by m_n the maximum of $|d_n|$ on the cubes neighboring H in $\mathbb{D}_{n-1}$, we certainly have the estimate $|(M_{A_k} d_n - d_n)(x)| \leq m_n(x)$ for every $x \in H$. It follows, since m_n is also constant on H, that

$$\begin{aligned} \int_H |(M_{A_k} d_n - d_n)(x)|^2 &\leq c \cdot 2^{(d-1)n} \cdot 2^k \cdot \sup_{x \in H} m_n^2(x) \\ &= c \cdot 2^{k-n} \cdot \int_H m_n^2. \end{aligned}$$

Summing the above over $H \in \mathbb{D}_{n-1}$ and noting that $\int_{\mathbb{Z}^d} m_n^2 \leq c \cdot \int_{\mathbb{Z}^d} |d_n|^2$, we have finished the proof of (3.2).

We now turn to the proof of (3.1) in case $k \geq n$. In this case we have $\mathcal{E}_k d_n = 0$, so it is enough to prove

$$\|M_{A_k} d_n\|_2^2 \leq c \cdot 2^{n-k} \|d_n\|_2^2. \tag{3.3}$$

In fact, we will prove the following pointwise estimate:

$$|M_{A_k} d_n(x)|^2 \le c \cdot 2^{n-k} \cdot M_{H_k} |d_n(x)|^2. \tag{3.4}$$

We get (3.3) by integrating both sides of (3.4) on $\mathbb{Z}^d$.

The proof of (3.4) involves a computation we are going to use quite a few times later, so we formulate it in a lemma:

LEMMA 3.1: *There is a constant c so that for any A which is a set-theoretic difference of two sets from $\mathcal{A}_k$ and for any $B \in \mathcal{A}_k$ we have*

$$\left(\frac{1}{|B|} \sum_{A+x} d_n\right)^2 \le c \cdot 2^{n-k} \cdot \frac{1}{|H_k|} \sum_{A+x} |d_n|^2, \quad n \le k.$$

Note that for any $A \in \mathcal{A}_k$, one can find an $A' \in \mathcal{A}_k$ which is disjoint from it, hence $A \setminus A' = A$. This is because the side-length of any $A \in \mathcal{A}_k$ has side-length less than 2^k, which is half that of H_k.

The inequality in (3.4) follows by taking $A = B = A_k$ in the lemma.

Proof of Lemma 3.1: We know that for every $H \in \mathbb{D}_n$ we have $\int_H d_n = 0$. Hence we can estimate

$$\begin{aligned}\sum_{A+x} d_n &= \sum_{H \in \mathbb{D}_n} \sum_{(A+x)\cap H} d_n \\ &= \sum_{\substack{H \in \mathbb{D}_n \\ H \cap \partial(A+x) \neq \emptyset}} \sum_{(A+x)\cap H} d_n.\end{aligned}$$

Denoting $\mathcal{B}(A) = \mathcal{B}(A,n,x) = \bigcup\{(A+x)\cap H \mid H \in \mathbb{D}_n,\ H \cap \partial(A+x) \neq \emptyset\}$, we can then estimate, using Cauchy's inequality,

$$\left(\sum_{A+x} d_n\right)^2 \le |\mathcal{B}(A)| \cdot \sum_{\mathcal{B}(A)} |d_n|^2.$$

Now, since A is a difference of two sets belonging to $\mathcal{A}_k$, the measure of $\mathcal{B}(A)$ is not more than a constant multiple of $2^n \cdot 2^{(d-1)k}$. Also, $B \in \mathcal{A}_k$ implies that $|B| \ge c \cdot 2^{kd} \ge |H_k|$, and hence dividing both sides by $|B|^2$ we get

$$\begin{aligned}\left(\frac{1}{|B|} \sum_{A+x} d_n\right)^2 &\le c \cdot 2^{n-k} \cdot \frac{1}{|H_k|} \sum_{\mathcal{B}(A)} |d_n|^2 \\ &\le c \cdot 2^{n-k} \cdot \frac{1}{|H_k|} \sum_{A+x} |d_n|^2. \qquad \blacksquare\end{aligned}$$

Proof of Theorem B: The structure of the proof is very similar to that of Theorem A′.

Write $f = \sum_n d_n$, where $d_1 f = f - \mathcal{E}_1 f$ and $d_n = \mathcal{E}_{n-1} f - \mathcal{E}_n f$ for $n > 1$. Denoting

$$\mathcal{S}_k = \mathcal{S}_k(x) = \{i \mid A_{t_i}, A_{t_{i+1}} \in \mathcal{A}_k\}$$

this time, we take $\mathbb{S}_k$ defined by $\mathbb{S}_k g(x) = (\sum_{i\in\mathcal{S}_k} |(M_{A_{t_i}} - M_{A_{t_{i+1}}})g|^2)^{1/2}(x)$, $u_n = v_n = d_n$ and $\sigma(j) = c \cdot 2^{-|j|/2}$ in the almost orthogonality lemma. It is enough to prove

$$\int_{\mathbb{Z}^d} \sum_{i\in\mathcal{S}_k} |(M_{A_{t_i}} - M_{A_{t_{i+1}}})d_n|^2(x) \le c \cdot 2^{-|n-k|} \|d_n\|_{\ell^2}^2. \tag{3.5}$$

To prove (3.5), let us first assume $n > k$. Recalling that $\mathbb{D}_n$ denotes the set of all atoms of the nth dyadic σ algebra σ_n, write

$$\int_{\mathbb{Z}^d} \sum_{i\in\mathcal{S}_k} |(M_{A_{t_i}} - M_{A_{t_{i+1}}})d_n|^2 = \sum_{H\in\mathbb{D}_{n-1}} \int_H \sum_{i\in\mathcal{S}_k} |(M_{A_{t_i}} - M_{A_{t_{i+1}}})d_n|^2.$$

We want to estimate $\int_H \sum_{i\in\mathcal{S}_k} |(M_{A_{t_i}} - M_{A_{t_{i+1}}})d_n|^2$ for $H \in \mathbb{D}_{n-1}$. Since d_n is constant on H, $\sum_{i\in\mathcal{S}_k} |(M_{A_{t_i}} - M_{A_{t_{i+1}}})d_n|^2$ can be nonzero only if for some $i \in \mathcal{S}_k$ at least one of the cubes $A_{t_i} + x$ or $A_{t_{i+1}} + x$ intersects the boundary of H. But by the definition $\mathcal{S}_k$, A_{t_i} and $A_{t_{i+1}}$ are both contained in H_k. Hence the set of x's for which $\sum_{i\in\mathcal{S}_k} |(M_{A_{t_i}} - M_{A_{t_{i+1}}})d_n|^2$ might be nonzero is contained in the set of x for which $H_k + x$ intersects the boundary of H. The measure of this latter set is not more than a constant multiple of $2^{(d-1)n} \cdot 2^k$:

$$\begin{aligned}
\Big(\sum_{i\in\mathcal{S}_k} |(M_{A_{t_i}} - M_{A_{t_{i+1}}})d_n|^2\Big)^{1/2}(x) &\le \sum_{i\in\mathcal{S}_k} |(M_{A_{t_i}} - M_{A_{t_{i+1}}})d_n|(x) \\
&\le \sum_{i\in\mathcal{S}_k} \frac{1}{|A_{t_{i+1}}|} \sum_{(A_{t_{i+1}} \setminus A_{t_i})+x} |d_n| \\
&\quad + \left(\frac{1}{|A_{t_i}|} - \frac{1}{|A_{t_{i+1}}|}\right) \cdot \sum_{A_{t_i}+x} |d_n| \\
&\le c \cdot \frac{1}{|H_k|} \sum_{H_k+x} |d_n| + \frac{1}{|H_k|} \sum_{H_k+x} |d_n| \\
&\le c \cdot M_{H_k}|d_n|(x).
\end{aligned}$$

Denoting by m_n the maximum of $|d_n|$ on the cubes neighboring H in $\mathbb{D}_{n-1}$, we certainly have the estimate $M_{H_k}|d_n|(x) \le m_n(x)$ for every $x \in H$. It follows, since m_n is also constant on H, that

$$\begin{aligned}\int_H \sum_{i\in\mathcal{S}_k} |(M_{A_{t_i}} - M_{A_{t_{i+1}}})d_n|^2(x) &\le c\cdot 2^{(d-1)n}\cdot 2^k \cdot \sup_{x\in H} m_n^2(x)\\ &= c\cdot 2^{k-n}\cdot \int_H m_n^2.\end{aligned}$$

Summing the above over $H\in\mathbb{D}_{n-1}$ and noting that $\int_{\mathbb{Z}^d} m_n^2 \le c\cdot\int_{\mathbb{Z}^d}|d_n|^2$, we have finished the proof of (3.5) in case $n>k$.

Let us now prove (3.5) for the case $k\ge n$. It is enough to prove the pointwise estimate

$$\sum_{i\in\mathcal{S}_k} |(M_{A_{t_i}} - M_{A_{t_{i+1}}})d_n|^2(x) \le c\cdot 2^{n-k}\cdot M_{H_k} d_n^2(x). \tag{3.6}$$

Indeed, we get (3.5) by integrating both sides of (3.6) on $\mathbb{Z}^d$.

Write first

$$\begin{aligned}\sum_{i\in\mathcal{S}_k} |(M_{A_{t_i}} - M_{A_{t_{i+1}}})d_n|^2(x) \le& \sum_{i\in\mathcal{S}_k}\left(\frac{1}{|A_{t_{i+1}}|}\sum_{(A_{t_{i+1}}\setminus A_{t_i})+x} d_n\right)^2\\ &+\sum_{i\in\mathcal{S}_k}\left(\frac{1}{|A_{t_i}|}-\frac{1}{|A_{t_{i+1}}|}\right)^2\cdot\left|\sum_{A_{t_i}+x} d_n\right|^2.\end{aligned}$$

Taking $A=A_{t_{i+1}}\setminus A_{t_i}$ and $B=A_{t_{i+1}}$ in Lemma 3.1, we can estimate

$$\begin{aligned}\sum_{i\in\mathcal{S}_k}\left(\frac{1}{|A_{t_{i+1}}|}\sum_{(A_{t_{i+1}}\setminus A_{t_i})+x} d_n\right)^2 &\le c\cdot 2^{n-k}\cdot\frac{1}{|H_k|}\sum_{i\in\mathcal{S}_k}\sum_{(A_{t_{i+1}}\setminus A_{t_i})+x} d_n^2\\ &\le c\cdot 2^{n-k}\cdot\frac{1}{|H_k|}\sum_{H_k} d_n^2,\end{aligned}$$

where in the last step we used the fact that the sets A_{t_i} are nested for $i\in\mathcal{S}_k$.

In order to finish the proof of the inequality in (3.6), we just have to estimate $\sum_{i\in\mathcal{S}_k}(\frac{1}{|A_{t_i}|}-\frac{1}{|A_{t_{i+1}}|})^2\cdot|\sum_{A_{t_i}+x} d_n|^2$. But this is accomplished by again using Lemma 3.1 (with letting both A and B run independently over all A_{t_i}, $i\in\mathcal{S}_k$):

$$\begin{aligned}\sum_{i\in\mathcal{S}_k}\left(\frac{1}{|A_{t_i}|}-\frac{1}{|A_{t_{i+1}}|}\right)^2\cdot\left|\sum_{A_{t_i}+x} d_n\right|^2 &\le \sup_{i\in\mathcal{S}_k}\left|\sum_{A_{t_i}+x} d_n\right|^2\cdot\left(\sum_{i\in\mathcal{S}_k}\frac{1}{|A_{t_i}|}-\frac{1}{|A_{t_{i+1}}|}\right)^2\\ &\le \sup_{i\in\mathcal{S}_k}\left|\sum_{A_{t_i}+x} d_n\right|^2\cdot\sup_{i\in\mathcal{S}_k}\frac{1}{|A_{t_i}|^2}\\ &\le c\cdot 2^{n-k}\cdot\frac{1}{|H_k|}\sum_{H_k} d_n^2. \quad\blacksquare\end{aligned}$$

4. Proof of the weak $(1,1)$ inequalities

The proofs follow the classical scheme of Calderón–Zygmund, but we need to make a modification in estimating the "bad" part of a function. Namely, instead of the ℓ^1 norm, we will estimate the ℓ^2 norm of the bad function off the set where the maximal function is large. We made a similar modification to the Calderón–Zygmund scheme in [9].

Let us now give more details. Suppose we want to prove that an operator $\mathbb{O}$ is of weak type $(1,1)$, that is we want to prove the inequality

$$|\{\mathbb{O}f > \alpha\}| \leq c \cdot \frac{\|f\|_{\ell^1}}{\alpha}.$$

Since this inequality is homogeneous in α we can assume $\alpha = 1$, so we need to show

$$|\{\mathbb{O}f > 1\}| \leq c \cdot \|f\|_{\ell^1}.$$

First, we decompose f into good and bad parts: $f = g + b$. The properties of this decomposition are summarized in the following result of Calderón–Zygmund (cf. [6]).

LEMMA 4.1 (Calderón–Zygmund decomposition): *Any $f \in \ell^1(\mathbb{Z}^d)$ can be written in the form $f = g + b$ where*

(1) $\|g\|_2^2 \leq c \cdot \|f\|_1$;

(2) $b = \sum_j b_j$ *where each b_j satisfies:*

  (a) *for some n, b_j is supported in a cube $B_j \in \mathbb{D}_n$,*

  (b) $\int_{\mathbb{Z}^d} b_j = \int_{B_j} b_j = 0$,

  (c) $\|b_j\|_1 \leq c \cdot |B_j|$,

  (d) $\sum_j |B_j| \leq \|f\|_1$.

We note that the cubes B_j in the lemma are pairwise disjoint. In the rest of this section, B_j refers to a cube from the Calderón–Zygmund decomposition, and we denote by $\mathbb{B}$ the collection of all the B_j.

We now estimate as

$$|\{\mathbb{O}f > 1\}| \leq |\{\mathbb{O}g > 1/2\}| + |\{\mathbb{O}b > 1/2\}|.$$

The operators we consider here have been proved to be type $(2,2)$ in Section 3, which makes the estimation of the good part obvious by 1. above:

$$\begin{aligned} |\{\mathbb{O}g > 1/2\}| &\leq 4 \cdot \int |\mathbb{O}g|^2 \\ &\leq c \cdot \|g\|_2^2 \\ &\leq c \cdot \|f\|_1. \end{aligned}$$

Before estimating the bad part, let us introduce a notation. For each j let $\tilde{B}_j$ denote the cube with the same center as B_j, but with each dimension expanded three times. We can write

$$|\{\mathbb{O}b > 1/2\}| = |\cup_j \tilde{B}_j \bigcap \{\mathbb{O}b > 1/2\}| + |(\mathbb{Z}^d \setminus \cup_j \tilde{B}_j) \bigcap \{\mathbb{O}b > 1/2\}|.$$

The first term is estimated, using 2./(d), as

$$\begin{aligned} |\cup_j \tilde{B}_j \bigcap \{\mathbb{O}b > 1/2\}| &\le \sum_j |\tilde{B}_j| \\ &= 3^d \cdot \sum_j |B_j| \\ &\le 3^d \cdot \|f\|_1. \end{aligned}$$

The second term is estimated as

$$|(\mathbb{Z}^d \setminus \cup_j \tilde{B}_j) \bigcap \{\mathbb{O}b > 1/2\}| \le 4 \cdot \int_{\mathbb{Z}^d \setminus \cup_j \tilde{B}_j} |\mathbb{O}b|^2.$$

We see that $\mathbb{O}$ will be weak type $(1,1)$ if the integral $\int_{\mathbb{Z}^d \setminus \cup_j \tilde{B}_j} |\mathbb{O}b|^2$ is successfully estimated by a constant multiple of $\|f\|_1$. Indeed, in the proof of both Theorem A′ and Theorem B, we will prove a bound of the form

$$\int_{\mathbb{Z}^d \setminus \cup_j \tilde{B}_j} |\mathbb{O}b|^2 \le c \cdot \sum_j |B_j|,$$

which, by 2./(d), is a sufficient bound.

Proof of Theorem A′: By the preliminary remarks, it is enough to prove the inequality

$$\int_{\mathbb{Z}^d \setminus \cup_j \tilde{B}_j} \sum_n \sup_{A \in \mathcal{A}_k} |M_A b - \mathcal{E}_k b|^2 \le c \cdot \sum_j \|b_j\|_1.$$

We will prove this inequality using the Almost Orthogonality Lemma, Lemma 2.1. Let $A_k \in \mathcal{A}_k$ be such that $\sup_{A \in \mathcal{A}_k} |M_A b - \mathcal{E}_k b| = |M_{A_k} b - \mathcal{E}_k b|$. Let us group together the B_j of the same size, so write $b = \sum_n h_n$ where

$$h_n = \sum_{B_j \in \mathbb{D}_n \cap \mathbb{B}} b_j.$$

The fact that we just consider x which is in $\mathbb{Z}^d \setminus \cup_j \tilde{B}_j$ has two important consequences.

First, $\mathcal{E}_k b(x) = 0$ for every k. Indeed, if $n \ge k$ then $\mathcal{E}_k h_n(x) = 0$ because the atom of σ_k containing x is disjoint from the support of h_n. On the other hand, if $n < k$, for each B_j in the support of h_n, the atom of σ_k containing x is either disjoint from B_j or it contains the entire B_j. But then, by 2./(b), $\mathcal{E}_k b_j(x) = 0$ for all B_j in the support of h_n, hence $\mathcal{E}_k h_n(x) = 0$.

The second consequence of x being in $\mathbb{Z}^d \setminus \cup_j \tilde{B}_j$ is that $M_{A_k} h_n(x) = 0$ for $n \geq k$. This is because $x + A_k$ is disjoint from any of the B_j in the support of h_n.

As a consequence of this discussion, we see it is enough to prove

$$\int_{\mathbb{Z}^d \setminus \cup_j \tilde{B}_j} \sum_k \left| M_{A_k} \sum_{n<k} h_n \right|^2 \leq c \cdot \sum_j \|b_j\|_1.$$

Let

$$d_n = \sum_{B_j \in \mathbb{D}_n \cap \mathbb{B}} \mathbf{1}_{B_j}.$$

Since

$$\begin{aligned}\sum_n \|d_n\|_{\ell^2}^2 &= \sum_n \|d_n\|_{\ell^1} \\ &= \sum_j |B_j|,\end{aligned}$$

taking $\mathbb{S}_k$ defined by $\mathbb{S}_k g(x) = M_{A_k} g(x)$, $u_n = h_n$, $v_n = d_n$ and $\sigma(j) = c \cdot 2^{-|j|/2}$ in the almost orthogonality lemma, Lemma 2.1, we see it is enough to prove, for $k > n$,

$$\|M_{A_k} h_n\|_{\ell^2}^2 \leq c \cdot 2^{n-k} \cdot \|d_n\|_{\ell^2}^2.$$

This follows if we prove the pointwise inequality

$$|M_{A_k} h_n|^2 \leq c \cdot 2^{n-k} M_{2H_k} |d_n|. \tag{4.1}$$

Indeed, integrate both sides of the above and use the fact that $d_n = d_n^2$.

The proof of (4.1) involves a computation we will repeat several times later, so we formulate it in a lemma that is similar to Lemma 3.1:

LEMMA 4.2: *There is a constant c so that for any A which is a set-theoretic difference of two sets from $\mathcal{A}_k$ and any $B \in \mathcal{A}_k$ we have the two pointwise estimates*

$$\begin{aligned}\left(\frac{1}{|B|} \sum_{A+x} h_n\right)^2 &\leq c \cdot 2^{n-k} \cdot \frac{1}{|H_k|} \sum_{A+x} |h_n|, \quad n < k; \\ \sum_{A+x} |h_n| &\leq \sum_{2H_k+x} |d_n|, \quad n < k.\end{aligned}$$

The inequality in (4.1) follows from the lemma by taking $A = B = A_k$.

Proof of Lemma 4.2: The proof of the second estimate is simple since, by 2./(c),

$$\begin{aligned}
\sum_{A+x} |h_n| &\leq \sum_{H_k+x} |h_n| \\
&= \sum_{B_j \in \mathbb{D}_n \cap \mathbb{B}} \sum_{(H_k+x)\cap B_j} |h_n| \\
&\leq \sum_{\substack{B_j \in \mathbb{D}_n \cap \mathbb{B} \\ B_j \cap (H_k+x) \neq \emptyset}} \sum_{B_j} |h_n| \\
&\leq \sum_{\substack{B_j \in \mathbb{D}_n \cap \mathbb{B} \\ B_j \cap (H_k+x) \neq \emptyset}} \sum_{B_j} d_n \\
&\leq \sum_{2H_k+x} d_n.
\end{aligned}$$

Let us now prove the first inequality of the lemma. We certainly have

$$\left| \sum_{A+x} h_n \right| \leq \sum_{A+x} |h_n|.$$

On the other hand, by property 2./(b), we have

$$\begin{aligned}
\sum_{A} h_n &= \sum_{B_j \in \mathbb{D}_n \cap \mathbb{B}} \sum_{A \cap B_j} h_n \\
&= \sum_{\substack{B_j \in \mathbb{D}_n \cap \mathbb{B} \\ B_j \cap \partial A \neq \emptyset}} \sum_{A \cap B_j} h_n.
\end{aligned}$$

Hence, introducing $\mathcal{B}(A) = \mathcal{B}(A, n, x) = \bigcup\{B_j \mid B_j \in \mathbb{D}_n \cap \mathbb{B},\ B_j \cap \partial A \neq \emptyset\}$, we have, by 2./(c),

$$\begin{aligned}
\left| \sum_{A} h_n \right| &\leq \sum_{\mathcal{B}(A)} |h_n| \\
&\leq \sum_{\mathcal{B}(A)} |d_n| \\
&= |\mathcal{B}(A)|.
\end{aligned}$$

Now, since $A \in \mathcal{A}_k$ is a difference of two sets belonging to $\mathcal{A}_k$, the measure of $\mathcal{B}(A)$ is not more than a constant multiple of $2^n \cdot 2^{(d-1)k}$. Also, $B \in \mathcal{A}_k$ implies that $|B| \geq c \cdot 2^{kd} \geq |H_k|$, and hence we have, dividing both ends by $|B|$,

$$\left| \sum_{A} h_n \right| \leq c \cdot 2^{n-k}.$$

Putting the two estimates on $|\sum_A h_n|$ together, we obtain

$$\left(\frac{1}{|B|}\sum_A h_n\right)^2 \le c\cdot 2^{n-k}\cdot\frac{1}{|H_k|}\sum_{A+x}|h_n|. \qquad \blacksquare$$

Proof of Theorem B: The structure of the proof is very similar to that of the weak $(1,1)$ part of Theorem A′.

By the preliminary remarks in this section, it is enough to prove the inequality

$$\int_{\mathbb{Z}^d\setminus\bigcup_j\tilde{B}_j}\sum_k\sum_{i\in\mathcal{S}_k}|(M_{A_{t_i}}-M_{A_{t_{i+1}}})b|^2\le c\cdot\sum_j\|b_j\|_1.$$

Denote

$$h_n=\sum_{B_j\in\mathbb{D}_n}b_j \quad\text{and}\quad d_n=\sum_{B_j\in\mathbb{D}_n}\mathbf{1}_{B_j}.$$

By an argument similar to the one given in the proof of Theorem A′, it is enough to prove, for $n<k$, the pointwise estimate

$$\sum_{i\in\mathcal{S}_k}|(M_{A_{t_i}}-M_{A_{t_{i+1}}})h_n(x)|^2\le c\cdot 2^{n-k}\cdot M_{2H_k}|d_n|. \tag{4.2}$$

Write first

$$\begin{aligned}\sum_{i\in\mathcal{S}_k}|(M_{A_{t_i}}-M_{A_{t_{i+1}}})h_n|^2 &\le \sum_{i\in\mathcal{S}_k}\left(\frac{1}{|A_{t_{i+1}}|}\sum_{(A_{t_{i+1}}\setminus A_{t_i})+x}h_n\right)^2\\ &\quad+\sum_{i\in\mathcal{S}_k}\left(\frac{1}{|A_{t_i}|}-\frac{1}{|A_{t_{i+1}}|}\right)^2\cdot\left|\sum_{A_{t_i}+x}h_n\right|^2.\end{aligned}$$

Taking $A=A_{t_{i+1}}\setminus A_{t_i}$ and $B=A_{t_{i+1}}$ in the first inequality of Lemma 4.2, we can estimate

$$\begin{aligned}\sum_{i\in\mathcal{S}_k}\left(\frac{1}{|A_{t_{i+1}}|}\sum_{(A_{t_{i+1}}\setminus A_{t_i})+x}h_n\right)^2 &\le c\cdot 2^{n-k}\cdot\frac{1}{|H_k|}\sum_{i\in\mathcal{S}_k}\sum_{(A_{t_{i+1}}\setminus A_{t_i})+x}|h_n|\\ &\le c\cdot 2^{n-k}\cdot\frac{1}{|H_k|}\sum_{H_k+x}|h_n|\\ &\le c\cdot 2^{n-k}\cdot\frac{1}{|H_k|}\sum_{2H_k+x}|d_n|,\end{aligned}$$

where in the last step we used the second inequality of Lemma 4.2.

In order to finish the proof of the inequality in (4.2), we just have to estimate $\sum_{i\in\mathcal{S}_k}(\frac{1}{|A_{t_i}|}-\frac{1}{|A_{t_{i+1}}|})^2\cdot|\sum_{A_{t_i}+x}h_n|^2$. But this is accomplished by again using the inequalities in Lemma 4.2 (with letting both A and B run independently over all A_{t_i}, $i \in \mathcal{S}_k$):

$$\begin{aligned}\sum_{i\in\mathcal{S}_k}\left(\frac{1}{|A_{t_i}|}-\frac{1}{|A_{t_{i+1}}|}\right)^2\left|\sum_{A_{t_i}+x}h_n\right|^2 &\le \sup_{i\in\mathcal{S}_k}\left|\sum_{A_{t_i}+x}h_n\right|^2\left(\sum_{i\in\mathcal{S}_k}\frac{1}{|A_{t_i}|}-\frac{1}{|A_{t_{i+1}}|}\right)^2\\ &\le \sup_{i\in\mathcal{S}_k}\left|\sum_{A_{t_i}+x}h_n\right|^2\sup_{i\in\mathcal{S}_k}\frac{1}{|A_{t_i}|^2}\\ &\le c\cdot 2^{n-k}\cdot\frac{1}{|H_k|}\sum_{2H_k}|d_n|. \qquad \blacksquare\end{aligned}$$

5. Generalizations to other domains

Theorems 1.7–1.10 were stated for averages over cubes. But the reader should have no problem verifying that the proofs also work for averages over balls or, more generally, for averages over dilates of a convex region.

Before we try to indicate the possibility of further generalizations, let us point out that our methods do not work for averages over an *arbitrary* nested sequence of rectangles. Indeed, we pose the following problem:

PROBLEM 5.1: *Let $A_1 \subset A_2 \subset \cdots$ be a nested sequence of rectangles in $\mathbb{Z}^d$. Is it true that for every $f \in L^1(X,\Sigma,m,\tau)$ the square function*

$$\left(\sum_n |M_{A_n}f - M_{A_{n+1}}f|^2\right)^{1/2}$$

is finite almost everywhere?

In the above problem, it is important to require that $f \in L^1$; the case $f \in L^p$, $p > 1$, is settled in our earlier paper [10].

Describing the regions which could replace the cubes in Theorems 1.7–1.10 is equivalent to describing the regions in the sets $\mathcal{A}_k$.

The description of the sequence $(\mathcal{A}_k)$ starts with replacing the original definition of the cubes (H_k) by a more general sequence of rectangles:

Let $(s(k))$ be a sequence of $\mathbb{Z}_+^d$ vectors so that $s(k+1)$ is obtained from $s(k)$ by increasing at least one but maybe more coordinates by at least 1. We denote by H_k the rectangle

$$\prod_{1\le j\le d}[-2^{s_j(k)}, 2^{s_j(k)}) \subset \mathbb{Z}^d.$$

The most important property of the regions in $\mathcal{A}_k$ is going to be that they do not have a too wild boundary. This means that if $A \in \mathcal{A}_k$ and n is smaller than k, then the total volume of all the translates of H_n that intersect the boundary of A should be small compared to the volume of A (which will be a constant multiple of H_k). Indeed, what we need to describe is that this ratio of the volumes goes to 0 fast enough as the difference between n and k goes to ∞.

So let $(\sigma(j))_{j\in\mathbb{Z}}$ be a sequence of positive numbers with $\sum_j \sigma(j) < \infty$. For two regions $A, B \subset \mathbb{Z}^d$, we denote

$$\mathcal{B}(A,B) = \{x \mid \partial A \cap (B+x) \neq \emptyset\},$$

where recall that for a region $A \subset \mathbb{Z}^d$ we denote by ∂A the **boundary** of A, that is the set of points which have neighbors both from A and from the complement of A.

We are now ready to describe the properties of the regions in $\mathcal{A}_k$; they should satisfy:

(1) $A \subset H_k$ for $A \in \mathcal{A}_k$;
(2) there is a positive constant c so that $c|H_k| \leq |A|$ for every $A \in \mathcal{A}_k$ and k;
(3) for $n \leq k$, and for $A \in \mathcal{A}_k$ we have $|\mathcal{B}(A, H_n)|/|H_k| \leq \sigma^2(n-k)$.

Once $(\mathcal{A}_k)$ is defined, the concept of a regular sequence is the same as before: we call the sequence (A_t) of regions **regular** (with respect to $(\mathcal{A}_k)$) iff they satisfy the following conditions:

- for every n and $t \in [2^{n-1}, 2^n)$, we have $A_t \in \mathcal{A}_n$;
- $|A_s| \leq |A_t|$ if $s \leq t$;
- if $2^{n-1} \leq s \leq t < 2^n$ for some n, then $A_s \subset A_t$.

The last step in our general setup is to identify the reverse martingale. Denote the set of rectangles of the form

$$\prod_{1\leq j\leq d} [m_j 2^{s_j(k)}, (m_j+1)2^{s_j(k)}) \subset \mathbb{Z}^d$$

by $\mathbb{D}_k$. For a function $f : \mathbb{Z}^d \to \mathbb{R}$, we denote by $\mathcal{E}_k f = f_k$ the expectation of f with respect to the σ algebra generated by $\mathbb{D}_k$.

Now the reader should have no problems proving the appropriate generalizations of Theorems A′ and B, and hence the generalizations of Theorems 1.7–1.10.

In the rest of this section, we give examples for various choices of $\mathcal{A}_k$ for which the above assumptions are applicable.

Examples: These examples are in $\mathbb{Z}^2$, but easily extended to $\mathbb{Z}^d$ for any d. In all cases one can take $\sigma(j) = c \cdot 2^{-j/2}$. Except in the last case, we leave it up to the reader to describe the appropriate (H_k).

- Let $\mathcal{A}_k$ contain all disks containing 0 and of radius between 2^{k-1} and 2^k.

- Let A be a convex region in the plane, and let D_tA be the t-th dilate of A. Let $\mathcal{A}_k$ contain all translates of D_tA, $2^{k-1} \le t < 2^k$, which contain the origin.
- For a region A, let X_sA be the horizontal dilation and Y_sA be the vertical dilation by the factor s. Let A be a convex region in the plane, and let $\mathcal{A}_k$ contain all translates of $X_{2^u}Y_{4^u}A$, $k-1 \le u < k$, which contain the origin.
- Let A be the set-theoretic difference of two convex regions in the plane, and let $\mathcal{A}_k$ contain all translates of $X_{2^u}Y_{4^u}A$, $k-1 \le u < k$, which are contained in the rectangle $H_k = [-2^{k+1}, 2^{k+1}) \times [-4^{k+1}, 4^{k+1})$.

6. Analogs in $\mathbb{R}^d$

It is rather straightforward to extend our results to averages over regions in the Euclidean space $\mathbb{R}^d$. Instead of repeating almost verbatim the previous section, we just concentrate on a single, but general enough example in $\mathbb{R}^d$ to illustrate the point.

For a region A in $\mathbb{R}^d$, let X_s^jA be the dilation by the factor s in the jth coordinate direction, $j = 1, 2 \ldots, d$. Let A be a region in $\mathbb{R}^d$ with nonempty interior and with piecewise smooth boundary. Let $\alpha_1, \alpha_2, \ldots, \alpha_d$ be positive integer powers of 2. For each integer k (so we consider negative k as well!), let $\mathcal{A}_k$ contain those translates of $X^1_{\alpha_1^u} \ldots X^d_{\alpha_d^u}A$, $k-1 \le u < k$, which are contained in the rectangle $H_k = [-\alpha_1^{k+1}, \alpha_1^{k+1}) \times \cdots \times [-\alpha_d^{k+1}, \alpha_d^{k+1})$.

The collection $(A_t)_{-\infty<t<\infty} = (A_t)_{-\infty<t<\infty}(f, x)$ of regions is called **regular** iff:

- for every integer n and real number $t \in [n-1, n)$, we have $A_t \in \mathcal{A}_n$;
- $|A_s| \le |A_t|$ if $s \le t$;
- if $n-1 \le s \le t < n$ for some integer n, then $A_s \subset A_t$.

For each n, the nth σ algebra σ_n is generated by atoms of the form $[s_1\alpha_1^n, (s_1+1)\alpha_1^n) \times \cdots \times [s_d\alpha_d^n, (s_d+1)\alpha_d^n)$. We denote by $\mathcal{E}_n$ the expectation with respect to σ_n. The average $M_Af(x)$ of a function f: $\mathbb{R}^2 \to \mathbb{R}$ on the region $A \subset \mathbb{R}^2$ is defined as $M_Af(x) = \frac{1}{|A|}\int_A f(x+y)dy$.

With very little modifications to the proofs, the reader can prove the "continuous" ($\mathbb{R}^d$) analogs of Theorems A′, B and 1.7–1.10. Additional care has to be taken only in the selection of the regular sets $(A_t)_{t>0}$ for each x and f to make the operators Lebesgue measurable.

The easiest way to assure the measurability of the operators is to choose the A_t independently of x and f. For example, let $A_t = X^1_{\alpha_1^t} \cdots X^d_{\alpha_d^t}A$. In this case, the corresponding λ-jump, variational and oscillation operators are measurable for every locally integrable function. Indeed, the measurability of the operators is clear if we restrict the allowable set of t's to a finite set. By the continuity of the averages over the A_t, we then get measurability for the operators when the t's are restricted to a compact interval. Finally, the measurabilty of the operators allowing the full range of t's follows by taking a sequenece of intervals that exhaust the interval $(0,\infty)$.

We can now state Theorems C, D and 6.1–6.4 which are "continuous" analogs of Theorems A′, B and 1.7–1.10, respectively.

In these results, we assume that $A_t = X^1_{\alpha_1^t} \cdots X^d_{\alpha_d^t} A$.

Theorem C: *We have*

$$\left\| \left(\sum_n \sup_{A\in\mathcal{A}_n} |M_A f - \mathcal{E}_n f|^2 \right)^{1/2} \right\|_{L^p} \leq c_p \cdot \|f\|_{L^p} \quad (1 < p \leq 2),$$

and

$$\left| \left\{ \left(\sum_n \sup_{A\in\mathcal{A}_n} |M_A f - \mathcal{E}_n f|^2 \right)^{1/2} > \alpha \right\} \right| \leq \frac{c}{\alpha} \cdot \|f\|_{L^1}.$$

Recall the remark after Theorem A: the constants c, c_p here and elsewhere in the paper may depend on the dimension d.

Corollary: *We have*

$$\left\| \left(\sum_n \sup_{A,B\in\mathcal{A}_n} |M_A f - M_B f|^2 \right)^{1/2} \right\|_{L^p} \leq c_p \cdot \|f\|_{L^p} \quad (1 < p \leq 2),$$

and

$$\left| \left\{ \left(\sum_n \sup_{A,B\in\mathcal{A}_n} |M_A f - M_B f|^2 \right)^{1/2} > \alpha \right\} \right| \leq \frac{c}{\alpha} \cdot \|f\|_{L^1}.$$

Theorem D: *We have*

$$\left\| \sup_{(t_i)} \left(\sum_n \sum_{\substack{i \\ n-1\leq t_i<t_{i+1}<n}} |M_{A_{t_i}} f - M_{A_{t_{i+1}}} f|^2 \right)^{1/2} \right\|_{L^p} \leq c_p \cdot \|f\|_{L^p} \quad (1 < p \leq 2),$$

and

$$\left| \left\{ \sup_{(t_i)} \left(\sum_n \sum_{\substack{i \\ n-1\leq t_i<t_{i+1}<n}} |M_{A_{t_i}} f - M_{A_{t_{i+1}}} f|^2 \right)^{1/2} > \alpha \right\} \right| \leq \frac{c}{\alpha} \cdot \|f\|_{L^1}.$$

In the above, the supremum is taken over all increasing sequences $(t_i)_{i\in\mathbb{Z}}$.

THEOREM 6.1: *Let $(t_k)_{k\in\mathbb{Z}}$ be an increasing sequence of numbers. Then we have the inequalities*

$$\left\|\left(\sum_k |M_{A_{t_k}}f - M_{A_{t_{k+1}}}f|^2\right)^{1/2}\right\|_{L^p} \le c_p \cdot \|f\|_{L^p} \quad (1 < p \le 2),$$

and

$$\left|\left\{\left(\sum_k |M_{A_{t_k}}f - M_{A_{t_{k+1}}}f|^2\right)^{1/2} > \alpha\right\}\right| \le \frac{c}{\alpha}\cdot \|f\|_{L^1}.$$

The constants do not depend on the sequence (t_k).

THEOREM 6.2: *Let $(v_k)_{k\in\mathbb{Z}}$ be an increasing sequence of positive numbers. Then we have the inequalities*

$$\left\|\left(\sum_k \sup_{v_k\le s\le t<v_{k+1}} |M_{A_s}f - M_{A_t}f|^2\right)^{1/2}\right\|_{L^p} \le c_p \cdot \|f\|_{L^p} \quad (1 < p \le 2),$$

and

$$\left|\left\{\left(\sum_k \sup_{v_k\le s\le t<v_{k+1}} |M_{A_s}f - M_{A_t}f|^2\right)^{1/2} > \alpha\right\}\right| \le \frac{c}{\alpha}\cdot \|f\|_{L^1}.$$

The constants do not depend on the sequence (v_k).

THEOREM 6.3:

$$\|(J((M_{A_t}),\lambda))^{1/2}\|_{L^p} \le c_p \cdot \|f\|_{L^p} \quad (1 < p \le 2),$$

and

$$|\{\lambda\cdot (J((M_{A_t}),\lambda))^{1/2} > \alpha\}| \le \frac{c}{\alpha}\cdot \|f\|_{L^1}.$$

COROLLARY:

$$|\{J((M_{A_t}),\lambda) > N\}| \le \frac{c}{\lambda N^{1/2}}\cdot \|f\|_{L^1}.$$

THEOREM 6.4: *For every $\varrho > 2$, we have*

$$\left\|\left(\sup_{(t_i)}\sum_i |M_{A_{t_i}}f - M_{A_{t_{i+1}}}f|^\varrho\right)^{1/\varrho}\right\|_{L^p} \le c_{p,\varrho} \cdot \|f\|_{L^p} \quad (1 < p \le 2),$$

and

$$\left|\left\{\left(\sup_{(t_i)}\sum_i |M_{A_{t_i}}f - M_{A_{t_{i+1}}}f|^\varrho\right)^{1/\varrho} > \alpha\right\}\right| \le \frac{c_\varrho}{\alpha}\cdot \|f\|_{L^1}.$$

Above, the supremum is taken over all increasing sequences $(t_i)_{i\in\mathbb{Z}}$.

7. Open problems

PROBLEM 7.1: *Let $A_1 \subset A_2 \subset \cdots$ be a nested sequence of disks in $\mathbb{Z}^d$, and let $p > 2$.*

Is it true that then for every $f \in L^p(X, \Sigma, m, \tau)$, the square-function

$$\left(\sum_n |M_{A_n} f - M_{A_{n+1}} f|^2\right)^{1/2}$$

is finite almost everywhere?

PROBLEM 7.2: *Let $A_1 \subset A_2 \subset \cdots$ be a nested sequence of ellipsoids in $\mathbb{Z}^d$, and let $p \geq 1$.*

Is it true that then for every $f \in L^p(X, \Sigma, m, \tau)$, the square-function

$$\left(\sum_n |M_{A_n} f - M_{A_{n+1}} f|^2\right)^{1/2}$$

is finite almost everywhere?

PROBLEM 7.3: *Let $A_1 \subset A_2 \subset \cdots$ be a nested sequence of rectangles in $\mathbb{Z}^d$.*

Is it true that for every $f \in L^1(X, \Sigma, m, \tau)$, we have

$$|\{\lambda \cdot (J((M_{A_t}), \lambda))^{1/2} > \alpha\}| \leq \frac{c}{\alpha} \cdot \|f\|_{\ell^1}?$$

PROBLEM 7.4: *Let $A_1 \subset A_2 \subset \cdots$ be a nested sequence of cubes in $\mathbb{Z}^d$.*

Is it true that for every $f \in L^1(X, \Sigma, m, \tau)$, we have

$$|\{J((M_{A_t}), \lambda) > N\}| \leq \frac{c}{\lambda \cdot N} \cdot \|f\|_{\ell^1}?$$

Finally, for the reader's convenience we repeat

PROBLEM 7.5: *Let $A_1 \subset A_2 \subset \cdots$ be a nested sequence of rectangles in $\mathbb{Z}^d$.*

Is it true that for every $f \in L^1(X, \Sigma, m, \tau)$, the square-function

$$\left(\sum_n |M_{A_n} f - M_{A_{n+1}} f|^2\right)^{1/2}$$

is finite almost everywhere?

8. Notes

The transference principle appears in [5].

Oscillational inequalities for the usual ergodic averages were first considered by Gaposhkin in [7, 8]. He considered the case $n_k = 2^k$. Then J. Bourgain considered oscillation inequalities in [3]. The inequality is rather hidden in [3, inequality (7.10)]. The method there works only for the subsequence $\frac{1}{2^k}\sum_{n\leq 2^k} f(\tau^n)$. H. White has some improvements on Bourgain's method; his work appears in his Master's Thesis at The University of North Carolina and was published in [1]. The method of [9] used to prove weak $(1,1)$ bounds for oscillation inequalities does not work for higher dimensional actions (this fact was the main motivation for the present paper).

Variational inequalities for the usual ergodic averages were first considered by J. Bourgain in [4, Corollary 3.25].

Upcrossing inequalities for the usual ergodic averages were proved by Bishop, [2]. Curiously, Bishop's methods work only for the averages $\frac{1}{t}\sum_{n\leq t} f(\tau^n)$, and they do not work for the symmetric averages, $\frac{1}{2t+1}\sum_{|n|\leq t} f(\tau^n)$, or for higher dimensional averages. Kalikow and Weiss in [11] devise a method to handle the symmetric averages as well as symmetric higher dimensional averages. Kalikow and Weiss's method does not seem to be able to handle arbitrary nested sequence of cubes. The methods in [9] handle the averages $\frac{1}{|I_N|}\sum_{n\in I_N} f(\tau^n)$ for any nested sequence (I_N) of intervals. The methods in [9] do not work for higher dimensional actions.

The Calderón–Zygmund decomposition is in [6].

The simple proof of the almost orthogonality lemma is due to E. Lesigne; he is the one who noticed that it follows immediately from the convolution inequality, $\|\sigma * b\|_{\ell^2} \leq \|\sigma\|_{\ell^1}\cdot\|b\|_{\ell^2}$.

ACKNOWLEDGEMENT: We thank the referee for an unusually thorough report, which contained several remarks that improved the presentation of the paper and cleared up some ambiguities in our work.

References

[1] I. Assani, K. Petersen and H. White, *Some connections between ergodic theory and harmonic analysis*, in *Almost Everywhere Convergence* (A. Bellow and R. L. Jones, eds.), Academic Press, San Diego, 1991, pp. 17–40.

[2] E. Bishop, *Foundations of Constructive Analysis*, McGraw-Hill, New York, 1967.

[3] J. Bourgain, *On the maximal ergodic theorem for certain subsets of the integers*, Israel Journal of Mathematics **61** (1988), 39–72.

[4] J. Bourgain, *Pointwise ergodic theorems for arithmetic sets (Appendix: The return time theorem)*, Publications Mathématiques de l'Institut des Hautes Études Scientifiques **69** (1989), 5–45.

[5] A. P. Calderón, *Ergodic theory and translation invariant operators*, Proceedings of the National Academy of Sciences of the United States of America **59** (1968), 349–353.

[6] A. P. Calderón and A. Zygmund, *On the existence of certain singular integrals*, Acta Mathematica **88** (1952), 85–139.

[7] V. F. Gaposhkin, *A theorem on the convergence almost everywhere of a sequence of measurable functions, and its applications to sequences of stochastic integrals*, Mathematics of the USSR-Sbornik **33** (1977), 1–17.

[8] V. F. Gaposhkin, *Individual ergodic theorems for normal operators on L^2*, Functional Analysis and its Applications **15** (1981), 14–18.

[9] R. Jones, R. Kaufman, J. Rosenblatt and M. Wierdl, *Oscillation in ergodic theory*, Ergodic Theory and Dynamical Systems **18** (1998), 889–935.

[10] R. Jones, J. Rosenblatt and M. Wierdl, *Oscillation inequalities for rectangles*, Proceedings of the American Mathematical Society **129** (2001), 1349–1358.

[11] B. Weiss and S. Kalikow, *Fluctuations of ergodic averages*, Illinois Journal of Mathematics **43** (1999), 480–488.