Ergod. Th. & Dynam. Sys. (1998), **18**, 889–935
Printed in the United Kingdom © *1998 Cambridge University Press*

Oscillation in ergodic theory

ROGER L. JONES†¶, ROBERT KAUFMAN‡, JOSEPH M. ROSENBLATT‡‖ and MÁTÉ WIERDL§∗∗

† *Department of Mathematics, DePaul University, 2219 N. Kenmore, Chicago, IL 60614, USA*
(e-mail: rjones@condor.depaul.edu)
‡ *Department of Mathematics, University of Illinois at Urbana, Urbana, IL 61801, USA*
(e-mail: jrsnblltt@symcom.math.uiuc.edu)
§ *Department of Mathematical Sciences, University of Memphis, Memphis, TN 38152, USA*
(e-mail: wierdlm@mathsci.msci.memphis.edu)

(*Received* 8 *October* 1996 *and accepted in revised form* 18 *March* 1997)

Abstract. In this paper we establish a variety of square function inequalities and study other operators which measure the oscillation of a sequence of ergodic averages. These results imply the pointwise ergodic theorem and give additional information such as control of the number of upcrossings of the ergodic averages. Related results for differentiation and for the connection between differentiation operators and the dyadic martingale are also established.

1. *Introduction*

In this paper we consider square functions in ergodic theory, martingale theory and differentiation. We also introduce square functions that show how the averages from these areas are related. Here is a brief description of each section.

Section 1. Here we introduce some concepts used throughout the paper, and state some typical results.

Section 2. Here we deal with inequalities in L^2. First a fundamental connection between differentiation and martingales is established by proving the L^2-boundedness of a square function involving these operators. Then we use these results to prove the L^2-boundedness for a great variety of ergodic square functions.

Section 3. We prove that the square functions of the previous section are weak type $(1, 1)$.

Section 4. Most of our square functions map L^∞ to BMO, and are type (p, p) for $1 < p < \infty$.

¶ R. Jones is partially supported by NSF Grant DMS—9302012.
‖ J. Rosenblatt is partially supported by NSF Grant DMS—9401093.
∗∗ M. Wierdl is partially supported by NSF Grant DMS—9500577.

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

Section 5. Some of the results on square functions lead to upcrossing results for ergodic averages. In particular, the exponential estimate of Kalikow and Weiss is generalized to averages taken over an arbitrary nested sequence of intervals.

Section 6. Here we summarize (and prove) results on martingales which are used in the paper or which motivated some of our results.

Before we can get into more details we need some definitions.

Definitions 1.1. For a sequence (x_n) and $\varrho \geq 1$ define the ϱth norm operator $\mathcal{S}_\varrho$ by

$$\mathcal{S}_\varrho(x_n) = \left(\sum_{n=1}^{\infty} |x_n|^\varrho \right)^{1/\varrho}.$$

For a fixed sequence $n_1 < n_2 < \cdots$, define the oscillational norm operator $\mathcal{O}$ by

$$\begin{aligned} \mathcal{O}_\varrho(x_n) &= \mathcal{S}_\varrho \left(\sup_{n_{k-1} \leq n \leq m < n_k} |x_n - x_m| \right) \\ &= \left(\sum_{k=1}^{\infty} \sup_{n_{k-1} \leq n \leq m < n_k} |x_n - x_m|^\varrho \right)^{1/\varrho}, \end{aligned}$$

and the variational norm operator $\mathcal{V}$ by

$$\begin{aligned} \mathcal{V}_\varrho(x_n) &= \sup_{(n(k))} \mathcal{S}_\varrho(x_{n(k)} - x_{n(k+1)}) \\ &= \sup_{(n(k))} \left(\sum_k |x_{n(k)} - x_{n(k+1)}|^\varrho \right)^{1/\varrho}, \end{aligned}$$

where the supremum is taken over all systems of indices $n(1) < n(2) < \cdots$.

Most of the time, we will use $\mathcal{S}_\varrho$ and $\mathcal{O}_\varrho$ when $\varrho = 2$. In this case we will often omit the subscript ϱ, that is, we write $\mathcal{S} = \mathcal{S}_2$, $\mathcal{O} = \mathcal{O}_2$.

Remark. It turns out that it is often more convenient to deal with the operator $\mathcal{O}'$ defined by

$$\mathcal{O}'(x_n) = \mathcal{S} \left(\sup_{n_{k-1} \leq n < n_k} |x_n - x_{n_{k-1}}| \right).$$

Let us note the following obvious inequalities:

$$\mathcal{O}'(x_n) \leq \mathcal{O}(x_n) \leq 2\mathcal{O}'(x_n), \tag{1.1}$$

and

$$\mathcal{O}(x_n) \leq 2\mathcal{S}(x_n). \tag{1.2}$$

Let us now describe some of our results. Let $R_1 \subset R_2 \subset \cdots$ be a nested sequence of intervals in $\mathbb{Z}$.

Warning. While the results of this paper apply to any nested sequence of intervals $R_1 \subset R_2 \subset \cdots$, from now on we assume, unless we say otherwise (as we often do in §5), that each R_n satisfies:

- $0 \in R_n$;
- $|R_n| = n$.

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

For the study of 'square function like' operators, these assumptions are without loss of generality: $0 \in R_n$ can be achieved by translating each interval by the same integer; $|R_n| = n$ can be achieved by inserting further intervals into the original sequence of intervals, and then re-indexing. These operations on the sequence $R_1 \subset R_2 \subset \cdots$ do not change our results.

Let (X, Σ, m, τ) denote a dynamical system with (X, Σ, m) a probability space and τ a measurable, invertible, measure-preserving point transformation from X to itself. For a function $f : X \to \mathbb{R}$, let

$$M_t f(x) = \frac{1}{t} \sum_{j \in R_t} f(\tau^j x).$$

For a sequence $n_1 < n_2 \cdots$ define the operator $O = O((n_k))$ by

$$\begin{aligned} Of(x) &= \mathcal{O}(M_t f(x)) \\ &= \left(\sum_{k=1}^{\infty} \sup_{n_{k-1} \le n \le m < n_k} |M_n f(x) - M_m f(x)|^2 \right)^{1/2}. \end{aligned} \tag{1.3}$$

The following theorem is a quantitative version of the pointwise ergodic theorem, but it generalizes several other results as well.

THEOREM A. *In any dynamical system the operator O defined in (1.3) is weak type $(1, 1)$, type (p, p) for $1 < p < \infty$ and maps L^∞ to BMO. The constants in the inequalities (although not necessarily the best constants) do not depend on the sequence $n_1 < n_2 < \cdots$.*

We prove that: O is type $(2, 2)$ in §2; O is weak type $(1, 1)$ in §3; O maps L^∞ to BMO in §4. That O is type (p, p) for $1 < p < \infty$ follows by interpolation.

There were many results that can be considered predecessors of Theorem A. They all dealt with the case $R_n = [0, n)$.

First, the ergodic square function S was considered; S is defined by

$$\begin{aligned} Sf(x) &= \mathcal{S}(M_{n_k} f(x) - M_{n_{k-1}} f(x)) \\ &= \left(\sum_{k=1}^{\infty} |M_{n_{k-1}} f(x) - M_{n_k} f(x)|^2 \right)^{1/2}. \end{aligned} \tag{1.4}$$

If $n_k = k$, S is the square function in [**19**] and for general (n_k) this is the square function considered in [**20**]. In [**20**], the operator S was shown to map L^2 to L^2, and L^1 to weak L^1, but that S is of type (p, p) for $2 < p < \infty$ is already new.

It was Gaposhkin [**15, 16**] (see also [**8**]) who first considered the operator O. He proved that O is type $(2, 2)$ if the sequence (n_k) satisfies $c \le n_{k+1}/n_k < C$ for some constants $1 < c < C$. White [**27**] (see [**2**]) modified techniques developed by Bourgain in [**7**] to prove that O is type $(2, 2)$ if the sequence (n_k) satisfies $n_{k+1} > c n_k^\beta$ for some $\beta > 1$.

The next theorem involves the operator V_ϱ defined by

$$\begin{aligned} V_\varrho f(x) &= \mathcal{V}_\varrho(M_t f(x)) \\ &= \sup_{(n(k))} \left(\sum_k |M_{n(k)} f(x) - M_{n(k+1)} f(x)|^\varrho \right)^{1/\varrho}. \end{aligned} \tag{1.5}$$

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

THEOREM B. *Let $\varrho > 2$. Then, in any dynamical system, the operator V_ϱ defined in (1.5) is weak type $(1, 1)$, type (p, p) for $1 < p < \infty$ and maps L^∞ to BMO.*

Note that if $\varrho = 2$, using the law of the iterated logarithm, it can be shown that for the Brownian martingale the operator V_ϱ is not type $(2, 2)$. This is exactly why we considered the operator O!

We prove that: V_ϱ is type $(2, 2)$ in §2; V_ϱ is weak type $(1, 1)$ in §3; V_ϱ maps L^∞ to BMO in §4. That V_ϱ is type (p, p) for $1 < p < \infty$ follows by interpolation.

The operator V_ϱ was first considered by Bourgain in [**6**]. He shows that if $R_n = \{0, 1, \ldots, (n-1)\}$ then V_ϱ is type $(2, 2)$.

We will use Theorem B in §5 to extend the exponential estimate of Kalikow and Weiss [**21**] for the number of upcrossings of the ergodic averages $M_t f(x)$.

As we mentioned earlier, many of our results are motivated by similar results for martingales. In fact, the results for martingales are *used* in many of the proofs. The fundamental tools in establishing the connection between martingales and ergodic averages are variants of the next theorem.

Let the differentiation operator D_n on $L^1([0, 1))$ be defined by

$$D_n f(x) = 2^n \int_x^{x+1/2^n} f(t)\, dt, \tag{1.6}$$

and let (f_k) denote the standard dyadic martingale, that is, $f_k = E(f|\mathcal{F}_k)$, where $\mathcal{F}_k$ is the σ-field generated by the intervals $[(j-1)/2^k, j/2^k)$, $j = 1, 2, \ldots, 2^k$. Let us define the square function S by

$$\begin{aligned} Sf(x) &= \mathcal{S}(D_k f(x) - f_k(x)) \\ &= \left(\sum_{k=1}^{\infty} |D_k f(x) - f_k(x)|^2 \right)^{1/2}. \end{aligned} \tag{1.7}$$

THEOREM C. *The operator S defined in (1.7) is weak type $(1, 1)$, type (p, p) for $1 < p < \infty$ and maps L^∞ to BMO.*

We prove that: S is type $(2, 2)$ in §2; S is weak type $(1, 1)$ in §3; S maps L^∞ to BMO in §4. That S is type (p, p) for $1 < p < \infty$ follows by interpolation.

Throughout the paper c and C (sometimes with additional parameters) will denote constants, but will not always be the same from one occurrence to the next. To avoid confusion, we will usually use ϕ to denote a function from $\mathbb{Z}$ to $\mathbb{R}$, or possibly from $\mathbb{Z}^d$ to $\mathbb{R}$. We will usually use f to denote a function from a probability space X to $\mathbb{R}$.

Remark 1.2. In general, one cannot conclude convergence from the boundedness of simple 'square function like' operators like S in (1.4). This is why we study more complicated operators which are associated with $\mathcal{O}$ and $\mathcal{V}$. In particular, let us show the existence of a sequence of operators, (T_n), and a function $f_0 \in L^2$, for which $(T_n f_0)$ fails to converge, but such that for any fixed sequence (n_k), the square function

$$Sf(x) = \left(\sum_{k=1}^{\infty} |T_{n_{k+1}} f(x) - T_{n_k} f(x)|^2 \right)^{1/2}$$

is a bounded operator on L^2.

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

Indeed, let Φ_j be the orthonormal system considered by Menschov (or we could let Φ_j be an appropriate rearrangement of any complete orthonormal system on $L^2[0,1)$) (see [**17**]). Then there is an $f_0 \in L^2[0,1)$ so that the partial sums $\sum_{j=1}^n c_j(f_0)\Phi_j(x)$ diverge a.e., where $c_k(f)$ is the Fourier coefficient of f with respect to Φ_k.

Let $T_k f(x) = \sum_{j=1}^k c_j(f)\Phi_j(x)$. Clearly, for any sequence (n_k) we will have

$$\left\| \left(\sum_{k=1}^{\infty} (T_{n_{k+1}} f(x) - T_{n_k} f(x))^2 \right)^{1/2} \right\|_2 \le \|f\|_2,$$

for all $f \in L^2[0,1)$. However, as we remarked above, $T_k f_0(x)$ fails to converge a.e.

2. *The L^2 theory*

Martingales and differentiation. We begin this section with a result relating martingales and the standard differentiation operator. Let $\mathcal{F}_k$ denote the dyadic σ-field on $[0,1)$ with 2^k atoms: $[0,1/2^k), [1/2^k, 2/2^k), \ldots, [(2^k-1)/2^k, 1)$. Thus $\mathcal{F}_0$ contains only the two elements $[0,1)$ and $\emptyset$. For $f \in L^2([0,1))$ let $f_k = E(f|\mathcal{F}_k)$. Define $d_0 = f_0 = E(f)$ and for $k > 0$ define $d_k = f_k - f_{k-1}$. Note that by orthogonality, $f = \sum_{k=0}^{\infty} d_k$ in norm. By the martingale convergence theorem $f = \sum_{k=0}^{\infty} d_k$ a.e., but this fact is not used in the following argument. For $f \in L^1([0,1))$, extended periodically to $\mathbb{R}$, define the operators

$$D_n f(x) = 2^n \int_x^{x+1/2^n} f(t)\,dt.$$

Consider the square function defined by

$$S(f)(x) = \mathcal{S}(D_n f(x) - f_n(x)) = \left(\sum_{n=1}^{\infty} |D_n f(x) - f_n(x)|^2 \right)^{1/2}. \tag{2.1}$$

THEOREM 2.1. *There is a constant C such that the square function in (2.1) satisfies*

$$\|Sf\|_2 \le C\|f\|_2.$$

Proof. We have

$$f_n - D_n f = E\left(\sum_{k=0}^{\infty} d_k \middle| \mathcal{F}_n \right) - D_n\left(\sum_{k=0}^{\infty} d_k \right) = \sum_{k=0}^{\infty} (E(d_k|\mathcal{F}_n) - D_n d_k).$$

Note that for $n \le k$ we have $E(f_k|\mathcal{F}_n) = f_n$. From this it follows that for $n < k$ we have $E(d_k|\mathcal{F}_n) = f_n - f_n = 0$. We also have $E(d_k|\mathcal{F}_n) = d_k$ for $k \le n$. Consequently

$$f_n - D_n f = \sum_{k=0}^{n} (d_k - D_n d_k) - \sum_{k=n+1}^{\infty} D_n d_k.$$

We claim that for $k \le n$ we have

$$\|d_k - D_n d_k\|_2^2 \le C 2^{k-n} \|d_k\|_2^2 \tag{2.2}$$

and that for $k > n$ we have

$$\|D_n d_k\|_2^2 \le C 2^{n-k} \|d_k\|_2^2. \tag{2.3}$$

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

If these two claims are true then by the triangle inequality and the above representation, we have

$$\|f_n - D_n f\|_2 \le \sum_{k=0}^{n} \|d_k - D_n d_k\|_2 + \sum_{k=n+1}^{\infty} \|D_n d_k\|_2,$$

and using (2.2) and (2.3), we have

$$\|f_n - D_n f\|_2 \le C \sum_{k=0}^{n} \sqrt{2}^{k-n} \|d_k\|_2 + C \sum_{k=n+1}^{\infty} \sqrt{2}^{n-k} \|d_k\|_2.$$

Consequently

$$\|f_n - D_n f\|_2^2 \le C\Big(\sum_{k=0}^{n} \sqrt{2}^{k-n} \|d_k\|_2 \Big)^2 + C\Big(\sum_{k=n+1}^{\infty} \sqrt{2}^{n-k} \|d_k\|_2 \Big)^2.$$

Writing $\sqrt{2} = 2^{1/4} \times 2^{1/4}$, and applying Hölder's inequality we have

$$\begin{aligned} \|f_n - D_n f\|_2^2 &\le C\Big(\Big(\sum_{k=0}^{n} \sqrt{2}^{k-n} \Big)^{1/2} \Big(\sum_{k=0}^{n} \sqrt{2}^{k-n} \|d_k\|_2^2 \Big)^{1/2} \Big)^2 \\ &\quad + C\Big(\Big(\sum_{k=n+1}^{\infty} \sqrt{2}^{n-k} \Big)^{1/2} \Big(\sum_{k=n+1}^{\infty} \sqrt{2}^{n-k} \|d_k\|_2^2 \Big)^{1/2} \Big)^2 \\ &\le C \sum_{k=0}^{n} \sqrt{2}^{k-n} \|d_k\|_2^2 + C \sum_{k=n+1}^{\infty} \sqrt{2}^{n-k} \|d_k\|_2^2. \end{aligned}$$

Summing over n we now have

$$\begin{aligned} \|Sf\|_2^2 &= \sum_{n=1}^{\infty} \|f_n - D_n f\|_2^2 \\ &\le C \sum_{n=0}^{\infty} \sum_{k=0}^{n} \sqrt{2}^{k-n} \|d_k\|_2^2 + C \sum_{n=0}^{\infty} \sum_{k=n+1}^{\infty} \sqrt{2}^{n-k} \|d_k\|_2^2 \\ &\le C \sum_{k=0}^{\infty} \|d_k\|_2^2 \sum_{n=k}^{\infty} \sqrt{2}^{k-n} + C \sum_{k=0}^{\infty} \|d_k\|_2^2 \sum_{n=0}^{k-1} \sqrt{2}^{n-k} \\ &\le C \sum_{k=0}^{\infty} \|d_k\|_2^2 = C \|f\|_2^2, \end{aligned}$$

where the last line follows by orthogonality.

It now remains to prove (2.2) and (2.3). We first prove (2.2). Fix x so that $r/2^k \le x < (r+1)/2^k$. Note that $d_k = f_k - f_{k-1}$ is constant on $[r/2^k, (r+1)/2^k)$. So if $x < (r+1)/2^k - 1/2^n$ then $D_n d_k(x) = d_k(x)$.

For the remaining x we have $|D_n d_k(x)| \le \max\{|d_k(x)|, |d_k(x + 1/2^n)|\}$ since $n \ge k$. Consequently

$$|D_n d_k(x)| \le \max\left\{ \left| d_k\left(\frac{r}{2^k}\right) \right|, \left| d_k\left(\frac{r+1}{2^k}\right) \right| \right\}.$$

From this we obtain

$$\|d_k - D_n d_k\|_2^2 \le \sum_{r=0}^{2^k-1} \int_{r/2^k}^{(r+1)/2^k} |d_k(x) - D_n d_k(x)|^2$$

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

$$\begin{aligned}
&= \sum_{r=0}^{2^k-1} \int_{(r+1)/2^k-1/2^n}^{(r+1)/2^k} |d_k(x) - D_n d_k(x)|^2 \\
&\le C \sum_{r=0}^{2^k-1} \frac{1}{2^n} \left(\left| d_k\left(\frac{r}{2^k}\right)\right|^2 + \left| d_k\left(\frac{r+1}{2^k}\right)\right|^2 \right) \\
&\le C 2^{k-n} \frac{1}{2^k} \sum_{r=0}^{2^k-1} \left| d_k\left(\frac{r}{2^k}\right)\right|^2 \\
&= C 2^{k-n} \|d_k\|_2^2.
\end{aligned}$$

To prove (2.3), let $k > n$. Fix x, $r/2^{k-1} \le x < (r+1)/2^{k-1}$. Then using the fact that $E(d_k|\mathcal{F}_{k-1}) = 0$, we have

$$\begin{aligned}
D_n d_k(x) &= 2^n \int_x^{x+1/2^n} d_k(t)\, dt \\
&= 2^n \left(\int_x^{(r+1)/2^{k-1}} d_k(t)\, dt + \int_{(r+2^{k-1-n})/2^{k-1}}^{x+1/2^n} d_k(t)\, dt \right).
\end{aligned}$$

Consequently,

$$\begin{aligned}
|D_n d_k(x)| \le\ & C 2^n \left(\left| d_k\left(\frac{r}{2^{k-1}}\right)\right| + \left| d_k\left(\frac{r}{2^{k-1}} + \frac{1}{2^k}\right)\right| \right. \\
& \left. + \left| d_k\left(\frac{r+2^{k-1-n}}{2^{k-1}}\right)\right| + \left| d_k\left(\frac{r+2^{k-1-n}}{2^{k-1}} + \frac{1}{2^k}\right)\right| \right) \frac{1}{2^k}.
\end{aligned}$$

From this we have

$$\begin{aligned}
\|D_n d_k\|_2^2 &\le \sum_{r=0}^{2^{k-1}-1} \int_{r/2^{k-1}}^{(r+1)/2^{k-1}} |D_n d_k(x)|^2\, dx \\
&\le C \sum_{r=0}^{2^k-1} \frac{1}{2^k} \left(2^{n-k} \left| d_k\left(\frac{r}{2^k}\right)\right| + 2^{n-k} \left| d_k\left(\frac{r+1}{2^k}\right)\right| \right)^2 \\
&= C (2^{n-k})^2 \sum_{r=0}^{2^k} \frac{1}{2^k} \left| d_k\left(\frac{r}{2^k}\right)\right|^2 \\
&= C 4^{n-k} \|d_k\|_2^2.
\end{aligned}$$

□

The above result can be generalized. For each positive integer n define a pair of measurable non-negative functions $a_n(x)$ and $b_n(x)$ such that $a_n(x) + b_n(x) = 1/2^n$. Let $D_n f(x) = 2^n \int_{x-a_n(x)}^{x+b_n(x)} f(t)\, dt$. Then we have the following theorem.

THEOREM 2.2. *There is a constant C, independent of the sequences $(a_n(x))$ and $(b_n(x))$, such that the square function defined by*

$$Sf(x) = \mathcal{S}(D_n f(x) - f_n(x)) = \left(\sum_{n=1}^{\infty} |D_n f(x) - f_n(x)|^2 \right)^{1/2}$$

satisfies

$$\|Sf\|_2 \le C \|f\|_2.$$

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

Proof. The proof is similar to the proof of Theorem 2.1. The only changes are in the proof that conditions (2.2) and (2.3) hold. For condition (2.2), the change is trivial because instead of bounding $|D_n d_k(x)|$ by $\max\{|d_k(r/2^k)|, |d_k((r+1)/2^k)|\}$, we have

$$|D_n d_k(x)| \le \max\left\{\left|d_k\left(\frac{r}{2^k}\right)\right|, \left|d_k\left(\frac{r+1}{2^k}\right)\right|, \left|d_k\left(\frac{r-1}{2^k}\right)\right|\right\}.$$

This only results in a change in the constant C that occurs in the statement of condition (2.2).

To see the analog of condition (2.3) we need to work harder. Here the most we can say is that

$$|D_n d_k(x)| \le C 2^n \frac{1}{2^k} \max\left\{\left|d_k\left(\frac{r}{2^{k-1}} + \frac{j}{2^k}\right)\right| : -2^{k-n} \le j \le 2^{k-n}\right\}.$$

Hence

$$\begin{aligned} |D_n d_k(x)|^2 &\le C 4^{n-k} \max\left\{\left|d_k\left(\frac{r}{2^{k-1}} + \frac{j}{2^k}\right)\right|^2 : -2^{k-n} \le j \le 2^{k-n}\right\} \\ &\le C 4^{n-k} \sum_{j=-2^{k-n}}^{2^{k-n}} \left|d_k\left(\frac{r}{2^{k-1}} + \frac{j}{2^k}\right)\right|^2. \end{aligned}$$

Using this we have

$$\begin{aligned} \|D_n d_k\|_2^2 &\le \sum_{r=0}^{2^{k-1}-1} \int_{r/2^{k-1}}^{(r+1)/2^{k-1}} |D_n d_k(x)|^2 \, dx \\ &\le C \sum_{r=0}^{2^k-1} \frac{1}{2^k} 4^{n-k} \sum_{j=-2^{k-n}}^{2^{k-n}} \left|d_n\left(\frac{r}{2^{k-1}} + \frac{j}{2^k}\right)\right|^2 \\ &\le C \sum_{r=0}^{2^k-1} \frac{1}{2^k} 4^{n-k} 2^{k-n+1} \left|d_n\left(\frac{r}{2^k}\right)\right|^2 \\ &\le C 2^{n-k} \sum_{r=0}^{2^k} \frac{1}{2^k} \left|d_k\left(\frac{r}{2^k}\right)\right|^2 \\ &= C 2^{n-k} \|d_k\|_2^2. \end{aligned}$$

□

We now have the following corollary.

COROLLARY 2.3. *For each n let $a_n(x), b_n(x), a'_n(x)$ and $b'_n(x)$ be four non-negative functions such that $a_n(x) + b_n(x) = a'_n(x) + b'_n(x) = 1/2^n$. If we let*

$$D_n f(x) = 2^n \int_{x-a_n(x)}^{x+b_n(x)} f(t)\, dt$$

and

$$\Delta_n f(x) = 2^n \int_{x-a'_n(x)}^{x+b'_n(x)} f(t)\, dt,$$

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

then the square function

$$Sf(x) = \mathcal{S}(D_n f(x) - \Delta_n f(x)) = \left(\sum_{n=1}^{\infty} |D_n f(x) - \Delta_n f(x)|^2 \right)^{1/2}$$

satisfies $\|Sf\|_2 \leq C\|f\|_2$, *where* C *is a universal constant and, in particular, does not depend on any of the sequences* a_n, b_n, a'_n *and* b'_n.

Proof. The proof follows from Theorem 2.2 and the pointwise estimate

$$\begin{aligned} Sf(x) &= \mathcal{S}(D_n f(x) - \Delta_n f(x)) \\ &\leq \mathcal{S}(D_n f(x) - f_n(x)) + \mathcal{S}(\Delta_n f(x) - f_n(x)). \end{aligned}$$

□

Consequences for ergodic averages. The previous results can be transferred to the integer setting, or the results can be reproved, using the same argument as above, and just changing the notation. Let $E_k\phi(r)$ denote conditional expectation on $\mathbb{Z}$ with respect to the dyadic σ-field. In particular, $E_k\phi(r) = (1/2^k)\sum_{t=j2^k}^{(j+1)2^k-1} \phi(t)$, where $j2^k \leq r < (j+1)2^k$. Let $R_1 \subset R_2 \subset \cdots$ denote a nested sequence of intervals in $\mathbb{Z}$ such that $|R_t| = t$ and $R_1 = \{0\}$. Let $M_t\phi(r) = (1/t)\sum_{j\in R_t} \phi(r+j)$. Then we have the following parallel result to Theorem 2.1.

THEOREM 2.4. *The square function*

$$\mathcal{S}(M_{2^n}\phi(r) - E_n\phi(r)) = \left(\sum_{n=1}^{\infty} |M_{2^n}\phi(r) - E_n\phi(r)|^2 \right)^{1/2}$$

maps $\ell^2(\mathbb{Z})$ *boundedly to* $\ell^2(\mathbb{Z})$.

Remark 2.5. In fact, this result also holds in the setting corresponding to Theorem 2.2, that is, if we replace the operators $M_{2^n} - E_n$ by the operators T_n defined by

$$T_n f(r) = \sup_R \left| \frac{1}{|R|} \sum_{j\in R} \phi(r+j) - E_n\phi(r) \right|,$$

where the supremum is taken over all intervals $R \subset \mathbb{Z}$ which contain 0, and of length 2^n.

The following result for dyadic averages on $\mathbb{Z}$ will be useful in the proof of some subsequent theorems.

THEOREM 2.6. *There is a universal constant* C *such that for any increasing sequence* (n_k) *we have*

$$\left\| \left(\sum_{k=1}^{\infty} \max_{n_{k-1}\leq n<n_k} |M_{2^n}\phi - M_{2^{n_k}}\phi|^2 \right)^{1/2} \right\|_{\ell^2} \leq C\|\phi\|_{\ell^2}. \tag{2.4}$$

Remark 2.7. Again, this result holds even if we replace the operators $M_{2^n} - M_{2^{n_k}}$ by the operators T_n defined by

$$T_n f(r) = \sup_{R,R'} \left| \frac{1}{|R|} \sum_{j\in R} \phi(r+j) - \frac{1}{|R'|} \sum_{j\in R'} \phi(r+j) \right|,$$

where the supremum is taken over all pairs of intervals $R, R' \subset \mathbb{Z}$ which contain 0, and $|R| = 2^n$, $|R'| = 2^{n_k}$.

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

Proof. First, note that

$$\left(\sum_{k=1}^{\infty} \max_{n_{k-1}\le n<n_k} |M_{2^n}\phi(x) - M_{2^{n_k}}\phi(x)|^2\right)^{1/2} \le \mathcal{O}(M_{2^n}\phi(x)),$$

and then estimate as

$$\mathcal{O}(M_{2^n}\phi(x)) \le \mathcal{O}(M_{2^n}\phi(x) - E_n\phi(x)) + \mathcal{O}(E_n\phi(x)).$$

The first term is estimated, by (1.2), as

$$\mathcal{O}(M_{2^n}\phi(x) - E_n\phi(x)) \le 2\mathcal{S}(M_{2^n}\phi(x) - E_n\phi(x)).$$

By Theorem 2.4, $\mathcal{S}(M_{2^n}\phi(x) - E_n\phi(x))$ is a bounded operator on $\ell^2(\mathbb{Z})$.

The ℓ^2-boundedness of the second term follows from Theorem 6.1 taking $f_n = E_n\phi$. □

Remark. Let (f_n) be a martingale. It follows from Theorem 6.1 that for any positive integers $p < q$, we have

$$\left\| \max_{p\le k<q} |f_k - f_p| \right\|_2 \le c\|f_p - f_q\|_2.$$

However, this is not true in the case of ergodic averages. Indeed, we will show the following. Let $p < q$ be even. Then in any aperiodic dynamical system, for any constant C, there is an $f \in L^2$ so that

$$\left\| \max_{p\le k<q} |M_k f - M_p f| \right\|_2 \ge C \max\left\{\|M_p f\|_2, \|M_q f\|_2\right\}.$$

For $\phi : \mathbb{Z} \to \mathbb{R}$ denote by $\|\cdot\|$ the 'density ℓ^2-norm':

$$\|\phi\| = \limsup_{L\to\infty} \sqrt{\frac{1}{2L} \sum_{-L\le j<L} \phi(j)^2}.$$

By the general principle described in [**1**, Theorem 2.3], it is sufficient to show that, for any constant C, there is $\phi : \mathbb{Z} \to \mathbb{R}$ with

$$\left\| \max_{p\le k<q} |M_k\phi - M_p\phi| \right\| \ge C \max\{\|M_p\phi\|, \|M_q\phi\|\}.$$

We show this in the following way: for some $\phi : \mathbb{Z} \to \mathbb{R}$ (for any even $p < q$)

$$\|M_{p+1}\phi - M_p\phi\| = \|M_{p+1}\phi\| = \frac{1}{p+1} \quad \text{and} \quad \|M_p\phi\| = \|M_q\phi\| = 0. \tag{2.5}$$

Indeed, let

$$\phi(x) = \begin{cases} 1 & \text{if } x \text{ even;} \\ -1 & \text{if } x \text{ odd.} \end{cases}$$

We clearly have, for every $x \in \mathbb{Z}$,

$$|M_t\phi(x)| = \begin{cases} 0 & \text{if } t \text{ even;} \\ 1/t & \text{if } t \text{ odd.} \end{cases}$$

from which (2.5) immediately follows.

Question 2.8. Is there a constant C so that, for every $p < q$,

$$\left\| \max_{p\le t<q} |M_t f - M_p f| \right\|_2 \le C \max_{p\le t<q} \|M_t f - M_q f\|_2.$$

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

Notation and preliminary results. To simplify some of the following discussion, we introduce some additional notation. For a sequence (x_n) let us denote

$$\|(x_n)\|_{v(\varrho)} = \sup\left(\sum_i |x_{n(i)} - x_{n(i+1)}|^\varrho\right)^{1/\varrho},$$

where the sup is taken over all systems of indices $n(1) < n(2) < \cdots$. That is, $\|(x_n)\|_{v(\varrho)} = V_\varrho(x_n)$. Note that for $\varrho \geq 1$, $\|\cdot\|_{v(\varrho)}$ is a semi-norm on the space of sequences. Also note that for $\varrho \geq 2$ we have the inequalities

$$\|(x_n)\|_{v(\varrho)} \leq \|(x_n)\|_{v(2)} \leq 2 \cdot \mathcal{S}(x_n), \tag{2.6}$$

where as before

$$\mathcal{S}(x_n) = \left(\sum_n |x_n|^2\right)^{1/2}.$$

Note that for every sequence (ϕ_n) of $\mathbb{Z} \to \mathbb{R}$ functions

$$\|\mathcal{S}(\phi_n)\|_{\ell^2} = \mathcal{S}(\|\phi_n\|_{\ell^2}),$$

that is to say, the operator $\mathcal{S}$ commutes with the ℓ^2-norm. We will repeatedly use the fact that for any $n_1 < n_2 < \cdots$ we have

$$\begin{aligned} \mathcal{S}(\|(x_n \mid n_{k-1} \leq n < n_k)\|_{v(2)}) &= \left(\sum_k \|(x_n \mid n_{k-1} \leq n < n_k)\|^2_{v(2)}\right)^{1/2} \\ &\leq \|(x_n)\|_{v(2)}, \end{aligned} \tag{2.7}$$

and if we also have $x_{n_k} = 0$ for all k, then

$$\|(x_n)\|_{v(2)} \leq 2 \cdot \mathcal{S}(\|(x_n \mid n_{k-1} \leq n < n_k)\|_{v(2)}). \tag{2.8}$$

It is easy to see that the constant 2 in (2.8) cannot be improved. Only the inequality in (2.8) needs proving.

Proof of (2.8). Let $n(1) < n(2) < \cdots$ be an arbitrary sequence. We need to show that

$$\sum_i |x_{n(i)} - x_{n(i+1)}|^2 \leq 4 \sum_k \|(x_n \mid n_{k-1} \leq n < n_k)\|^2_{v(2)}. \tag{2.9}$$

In the sum $\sum_i |x_{n(i)} - x_{n(i+1)}|^2$, the only terms which are not controlled by the sum on the right-hand side of (2.9) are those for which $n_{k-1} \leq n(i) < n_k$ and $n_{k'-1} \leq n(i+1) < n_{k'}$ with some $k < k'$. Let I denote the set of such 'bad' indices i. For $i \in I$, we write, using $x_{n_{k-1}} = x_{n_{k'-1}} = 0$,

$$\begin{aligned} &|x_{n(i)} - x_{n(i+1)}|^2 \\ &\quad\leq 2(|x_{n(i)}|^2 + |x_{n(i+1)}|^2) \\ &\quad= |x_{n(i)} - x_{n_{k-1}}|^2 + |x_{n(i)} - x_{n_{k-1}}|^2 + |x_{n(i+1)} - x_{n_{k'-1}}|^2 && \text{(2.10a)} \\ &\qquad + |x_{n(i+1)} - x_{n_{k'-1}}|^2. && \text{(2.10b)} \end{aligned}$$

Let S be the quantity we get by replacing, for each $i \in I$, the term $|x_{n(i)} - x_{n(i+1)}|^2$ by the term in (2.10b) in the sum $\sum_i |x_{n(i)} - x_{n(i+1)}|^2$. Then

$$S \leq \sum_k \|(x_n \mid n_{k-1} \leq n < n_k)\|^2_{v(2)}.$$

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

We have the same inequality if S denotes the sum over $i \in I$ of any of the terms in (2.10a), and this concludes our proof of (2.9). □

For a fixed sequence $n_1 < n_2 < \cdots$ define

$$\begin{aligned} \|(x_n)\|_{o(2)} &= \mathcal{O}(x_n) \\ &= \left(\sum_k \sup_{n_{k-1} \le n \le m < n_k} (x_n - x_m)^2 \right)^{1/2}. \end{aligned}$$

Note that $\|\cdot\|_{o(2)}$ is a semi–norm on the space of sequences and that the inequality

$$\|(x_n)\|_{o(2)} \le \|(x_n)\|_{v(2)}, \tag{2.11}$$

is valid for any fixed $n_1 < n_2 < \cdots$ in the definition of the norm $o(2)$.

Remark 2.9. It is possible for the right-hand side of inequality (2.11) to be much larger than the left-hand side. To see this we can consider the following simple example. Fix a sequence (n_k) and let $x_n = 1$ for $n_{k-1} \le n < n_k$ if k is even and $x_n = 0$ for $n_{k-1} \le n < n_k$ if k is odd. Clearly $\|(x_n)\|_{o(2)} = 0$ while $\|(x_n)\|_{v(2)} = \infty$. Consequently, results about the $v(2)$ norm are stronger than the associated result about the $o(2)$ norm.

As an illustration on the use of the above tools, we show how Gaposhkin's result—that the square function O defined in (1.3) is a bounded operator from L^2 to L^2 when $n_k = 2^k$—can be used to quickly obtain a result which strengthens not only Gaposhkin's own result but also improves White's result mentioned in the introduction.

PROPOSITION 2.10. *Suppose that the sequence $n_1 < n_2 < \cdots$ is lacunary. Then*

$$\| \|(M_t\phi)\|_{o(2)} \|_{\ell^2} \le C \cdot \|\phi\|_{\ell^2}.$$

Remark. Recall that Gaposkhin's result assumes the sequence (n_k) neither grows too fast nor too slowly. That is, he assumes there are constants $C > c > 1$ such that $c < n_{k+1}/n_k < C$. Here we are only assuming there is a constant c so that $1 < c < n_{k+1}/n_k$.

Proof. We can assume without loss of generality that $n_{k+1}/n_k > 2$. (If not, select a partition of the sequence into a finite number of subsequences for which this is true. Apply the following argument to each subsequence, and use the triangle inequality and the conclusion of the argument for the subsequences to obtain the result for the full sequence.) Let us define the operators V_t by

$$V_t f = M_{2^{k-1}}\phi; \quad 2^{k-1} \le t < 2^k.$$

Estimate as

$$\| \|(M_t\phi)\|_{o(2)} \|_{\ell^2} \le \| \|(V_t\phi)\|_{o(2)} \|_{\ell^2} + \| \|(M_t\phi - V_t\phi)\|_{o(2)} \|_{\ell^2}.$$

As for the first term, note that

$$\|(V_t\phi)\|_{o(2)} \le \|(M_{2^k}\phi)\|_{o(2)} + \mathcal{S}(M_{2^{k-1}}\phi - M_{2^k}\phi).$$

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

The reason we have to include the term $\mathcal{S}(M_{2^{k-1}}\phi - M_{2^k}\phi)$ is because we may have $2^{k-1} < n_j < 2^k$. Then $M_{2^{k-1}}$ may be in two terms of the sum given by $\|(V_t\phi)\|_{o(2)}$ while it could occur in only one term of the sum given by $\|(M_{2^k}\phi)\|_{o(2)}$. Note that $\|(M_{2^k}\phi)\|_{o(2)}$ is a bounded operator by Theorem 2.6 and $\mathcal{S}(M_{2^{k-1}}\phi - M_{2^k}\phi)$ is also a bounded operator from ℓ^2 to ℓ^2. A fast way to see this is to use Theorem 2.4 and the fact that the martingale square function is a bounded operator.

To bound the second term, define $K_t = M_t - V_t$. All we need to show is the pointwise estimate

$$\begin{aligned}\|(K_t\phi)\|_{o(2)} &\leq 4\cdot\mathcal{S}\left(\sup_{2^{j-1}\leq t<2^j}|K_t\phi|\right)\\ &= 4\cdot\mathcal{S}\left(\sup_{2^{j-1}\leq t<2^j}|M_t\phi - M_{2^j}\phi|\right)\end{aligned}$$

because then Gaposhkin's result shows that $\|(K_t\phi)\|_{o(2)}$ is a bounded $\ell^2\to\ell^2$ operator.

Thus our problem is reduced to showing a statement about scalar sequences. That is, it is enough to show that for every sequence (x_t) of real numbers, we have

$$\|(x_t)\|_{o(2)} \leq 4\cdot\mathcal{S}\left(\sup_{2^{j-1}\leq t<2^j}|x_t|\right).$$

Write

$$\|(x_t)\|_{o(2)}^2 = \sum_k (x_{t_k} - x_{s_k})^2 = (\mathcal{S}(x_{t_k} - x_{s_k}))^2,$$

with $n_{k-1} \leq s_k < t_k < n_k$. Since $n_{k+1}/n_k > 2$, there is at least 1 power of 2 between t_k and t_{k+2} and between s_k and s_{k+2}. But then, by the triangle inequality,

$$\begin{aligned}\mathcal{S}(x_{t_k} - x_{s_k}) &\leq S(x_{t_k}) + \mathcal{S}(x_{s_k})\\ &\leq 4\cdot\mathcal{S}\left(\sup_{2^{j-1}\leq t<2^j}|x_t|\right).\end{aligned}$$

□

By standard transference methods (see, for example, [**12, 25**]), Proposition 2.10 is easily generalized to any measure-preserving system. Assume τ is a measure-preserving point transformation (assume τ is invertible if any R_n contains negative integers) and define $M_n f(x) = (1/n)\sum_{k\in R_n} f(\tau^k x)$. Then we have the following corollary.

COROLLARY 2.11. *Suppose that the sequence $n_1 < n_2 < \cdots$ is lacunary. Then*

$$\|\|(M_t f)\|_{o(2)}\|_{L^2} \leq C\cdot\|f\|_{L^2}.$$

Inequalities for oscillation and variation. Gaposhkin's result can be generalized in a different direction as well. We will state and prove the result only in the special case where we use the dyadic blocks $[2^{k-1}, 2^k)$, but it is true with the more general lacunarity assumptions used by Gaposhkin.

PROPOSITION 2.12. *For all $\phi\in\ell^2$ we have*

$$\|\mathcal{S}(\|(M_t\phi \mid 2^{k-1}\leq t < 2^k)\|_{v(2)})\|_{\ell^2} \leq C\cdot\|\phi\|_{\ell^2}. \tag{2.12}$$

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

Further, for any dynamical system (X, Σ, μ, τ) *with* τ *an invertible measure-preserving point transformation, we have for any* $f \in L^2(X)$ *the inequality*

$$\|\mathcal{S}(\|(M_t f \mid 2^{k-1} \le t < 2^k)\|_{v(2)})\|_{L^2} \le C \cdot \|f\|_{L^2}.$$

Remark. The difference between this result and Gaposhkin's inequality is the use of the $v(2)$ norm in place of the $o(2)$ norm, where we take $n_k = 2^k$ in the definition of the $o(2)$-norm. Since, as we pointed out in Remark 2.9, the $v(2)$ norm can be much larger than the $o(2)$ norm, this is a stronger result.

Before proving Proposition 2.12 we show how it can be used to deduce two inequalities for the norms $\mathcal{V}_\varrho$, $\varrho > 2$, and $\mathcal{O}$.

THEOREM 2.13. *If* $2 < \varrho < \infty$ *then*

$$\|\|(M_t\phi)\|_{v(\varrho)}\|_{\ell^2} \le C(\varrho) \cdot \|\phi\|_{\ell^2}. \tag{2.13}$$

There is a constant C *so that for any sequence* $n_1 < n_2 < \cdots$ *we have*

$$\|\|(M_t\phi)\|_{o(2)}\|_{\ell^2} \le C \cdot \|\phi\|_{\ell^2}. \tag{2.14}$$

Further, if the averages M_t *are the ergodic averages*

$$M_t f(x) = \frac{1}{t} \sum_{k \in R_t} f(\tau^k x),$$

then

$$\|\|(M_t f)\|_{v(\varrho)}\|_{L^2} \le C(\varrho) \cdot \|f\|_{L^2}$$

and

$$\|\|(M_t f)\|_{o(2)}\|_{L^2} \le C \cdot \|f\|_{L^2}.$$

The inequality in (2.13) and its proof are slightly generalized versions of those of Bourgain, [**6**, Corollary 3.25]. The main difference is that we allow an arbitrary nested sequence of intervals $R_1 \subset R_2 \cdots$ in the definition M_t, while Bourgain proved his result only for $R_n = [0, n)$. Also, (2.14) extends Proposition 2.10 to any sequence (n_k).

Proof of the implication Proposition 2.12 $\Longrightarrow$ *Theorem 2.13.* Let the operators V_t be

$$V_t\phi = M_{2^{k-1}}\phi; \quad 2^{k-1} \le t < 2^k.$$

To prove inequality (2.13) we estimate

$$\|\|(M_t\phi)\|_{v(\varrho)}\|_{\ell^2} \le \|\|(V_t\phi)\|_{v(\varrho)}\|_{\ell^2} + \|\|(M_t\phi - V_t\phi)\|_{v(\varrho)}\|_{\ell^2}. \tag{2.15}$$

To handle the first term on the right-hand side of inequality (2.15) we again use the operators (E_k), where E_k is the conditional expectation with respect to the algebra generated by the intervals $[j \cdot 2^k, (j+1) \cdot 2^k)$, $j \in \mathbb{Z}$. Then we have

$$\|\|(V_t\phi)\|_{v(\varrho)}\|_{\ell^2} = \|\|(M_{2^k}\phi)\|_{v(\varrho)}\|_{\ell^2} \tag{2.16}$$

$$\le \|\|(E_k\phi)\|_{v(\varrho)}\|_{\ell^2} + \|\|(M_{2^k}\phi - E_k\phi)\|_{v(\varrho)}\|_{\ell^2}. \tag{2.17}$$

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

Here we know

$$\| \|(E_k\phi)\|_{v(\varrho)}\|_{\ell^2} \le C\|\phi\|_{\ell^2}$$

by Lepingle's result (see [**22, 23**]). For the last term in inequality (2.16) write first, using (2.6),

$$\| \|(M_{2^k}\phi - E_k\phi)\|_{v(\varrho)}\|_{\ell^2} \le 2 \cdot \|\mathcal{S}(M_{2^k}\phi - E_k\phi)\|_{\ell^2}.$$

But Theorem 2.4 says

$$\|\mathcal{S}(M_{2^k}\phi - E_k\phi)\|_{\ell^2} \le C\|\phi\|_{\ell^2}.$$

Hence inequality (2.16) is bounded by $C\|\phi\|_{\ell^2}$.

Bounding the second term on the right-hand side of inequality (2.15) is done as follows. Recall that

$$\|(M_t\phi - V_t\phi)\|_{v(\varrho)} \le \|(M_t\phi - V_t\phi)\|_{v(2)}.$$

By (2.8), we have

$$\begin{aligned} \|(M_t\phi - V_t\phi)\|_{v(2)} &\le 2 \cdot S(\|(M_t\phi - M_{2^k}\phi \mid 2^{k-1} \le t < 2^k)\|_{v(2)}) \\ &= 2 \cdot \mathcal{S}(\|(M_t\phi \mid 2^{k-1} \le t < 2^k)\|_{v(2)}). \end{aligned}$$

Assuming Proposition 2.12, we are now done with the proof of inequality (2.13), since Proposition 2.12 says the above expression is controlled by $\|\phi\|_{\ell^2}$.

We now prove inequality (2.14). Write, as before,

$$\| \|(M_t\phi)\|_{o(2)}\|_{\ell^2} \le \| \|(V_t\phi)\|_{o(2)}\|_{\ell^2} + \| \|(M_t\phi - V_t\phi)\|_{o(2)}\|_{\ell^2}. \tag{2.18}$$

As for the first term, first note

$$\|(V_t\phi)\|_{o(2)} \le \|(M_{2^k}\phi)\|_{o(2)} + \mathcal{S}(M_{2^{k-1}}\phi - M_{2^k}\phi).$$

Now we can apply Theorems 2.6 and 2.4 to show both these operators are bounded from ℓ^2 to ℓ^2.

The boundedness of the second term in (2.18) follows from the inequality

$$\|(M_t\phi - V_t\phi)\|_{o(2)} \le \|(M_t\phi - V_t\phi)\|_{v(2)}.$$

We saw in Proposition 2.12 above that this expression is controlled by $\|\phi\|_{\ell^2}$, and so the proof of (2.14) is complete.

The last statement, about the ergodic averages, follows from the prior statements and standard transfer methods. □

Remark. In the above proof we used Lepingle's theorem, that for martingales, and $s > 2$, we have

$$\| \|(E_k\phi)\|_{v(s)}\|_{\ell^2} \le C\|\phi\|_{\ell^2}.$$

We will show in §6 that for martingales, the operator $T\phi = \|(E_k\phi)\|_{v(s)}$ is weak type $(1, 1)$ and maps ℓ^p to ℓ^p for $1 < p < \infty$. We will see in the next two sections that the same is true for the associated operator $V_s\phi = \|(M_k\phi)\|_{v(s)}$.

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

Proof of Proposition 2.12. We will need estimates for the Fourier transform of M_t. We state the required estimate as a lemma so it will be easier to refer to later. Recall that

$$\widehat{M}_t(\beta) = \frac{1}{t}\sum_{n\in R_t} e(n\beta),$$

where $e(x) = e^{2\pi i x}$.

LEMMA 2.14. *Let $s < t$. Then for every $\beta \in [-1/2, 1/2]$ we have*

$$|\widehat{M}_t(\beta) - \widehat{M}_s(\beta)| \le C \cdot \min\left\{\frac{t-s}{t}, (t-s)|\beta|, \frac{1}{s|\beta|}\right\}.$$

Proof of Lemma 2.14. It is easy to see that

$$|\widehat{M}_t(\beta) - \widehat{M}_s(\beta)| \le \frac{C}{s|\beta|}.$$

Indeed, write

$$|\widehat{M}_t(\beta) - \widehat{M}_s(\beta)| \le |\widehat{M}_t(\beta)| + |\widehat{M}_s(\beta)|,$$

and use that for any interval R the numbers $e(n\beta)$, $n \in R$, form a geometric progression. Let us now show

$$|\widehat{M}_t(\beta) - \widehat{M}_s(\beta)| \le C \cdot \min\left\{\frac{t-s}{t}, (t-s)|\beta|\right\}.$$

This is the part of the proof where we are using the fact that we have a nested sequence of rectangles.

$$\begin{aligned}
|\widehat{M}_t(\beta) - \widehat{M}_s(\beta)| &= |\widehat{M}_t(\beta) - 1 - (\widehat{M}_s(\beta) - 1)| \\
&= \left|\frac{1}{t}\sum_{n\in R_t}(e(n\beta)-1) - \frac{1}{s}\sum_{n\in R_s}(e(n\beta)-1)\right| \\
&= \left|\left(\frac{1}{t}-\frac{1}{s}\right)\sum_{n\in R_s}(e(n\beta)-1) + \frac{1}{t}\sum_{n\in R_t\setminus R_s}(e(n\beta)-1)\right| \\
&\le C\cdot\left(\frac{1}{t}-\frac{1}{s}\right)\sum_{n\in R_s}|e(n\beta)-1| + \frac{1}{t}\sum_{n\in R_t\setminus R_s}|e(n\beta)-1|,
\end{aligned}$$

from which the estimate easily follows using that

$$|e(n\beta) - 1| \le C \cdot \min\{1, |n\beta|\}. \qquad \square$$

In fact we will use the following immediate consequence of the lemma.

COROLLARY 2.15. *Let $2^{k-1} \le t < 2^k$. Then for every $\beta \in [-1/2, 1/2]$ we have*

$$|\widehat{M}_t(\beta) - \widehat{M}_{2^k}(\beta)| \le C \cdot \min\left\{2^k|\beta|, \frac{1}{2^k|\beta|}\right\},$$

and

$$|\widehat{M}_t(\beta) - \widehat{M}_{t-1}(\beta)| \le \frac{C}{2^k}.$$

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

With the notation $K_t = M_t - M_{2^k}$, we can rewrite (2.12) as

$$\|S(\|(K_t\phi \mid 2^{k-1} \le t < 2^k)\|_{v(2)})\|_{\ell^2} \le C \cdot \|\phi\|_{\ell^2}. \tag{2.19}$$

We now write K_t as a sum of orthogonal elements by restricting the Fourier transform of K_t to disjoint sets, the union of which is the 1-periodic extension of $[-1/2, 1/2]\setminus\{0\}$. (It does not matter that 0 is excluded since $\widehat{K}(0) = 0$.)

First, for $j \in \mathbb{Z}$, we introduce the functions χ_j. For positive j, we let χ_j be the 1-periodic extension of the characteristic function of the set

$$[-2^{-j}, -2^{-j-1}) \cup [2^{-j-1}, 2^{-j}).$$

For $j \le 0$ we set $\chi_j \equiv 0$.

Next, for $2^{k-1} \le t < 2^k$, we define the operator $K_{j,t}$ by its Fourier transform as

$$\widehat{K}_{j,t}(\beta) = \widehat{K}_t(\beta)\chi_{j+k}(\beta).$$

Clearly,

$$K_t = \sum_{j\in\mathbb{Z}} K_{j,t}. \tag{2.20}$$

We need the following.

LEMMA 2.16.

- *For every j and t we have*
$$|\widehat{K}_{j,t}(\beta)| \le \frac{C}{2^{|j|}}. \tag{2.21}$$
- *For every j and t, if $2^{k-1} \le t < 2^k$, we have*
$$|\widehat{K}_{j,t}(\beta) - \widehat{K}_{j,t-1}(\beta)| \le \frac{C}{2^k}. \tag{2.22}$$

Proof of Lemma 2.16. We just need to prove (2.21); the inequality in (2.22) is an immediate consequence of the second estimate in Corollary 2.15.

Note that $\widehat{K}_{j,t}(\beta) \ne 0$ only if $2^{-j-k-1} \le |\beta| \le 2^{-j-k}$. From the inequalities of Corollary 2.15, we use

$$|\widehat{K}_t(\beta)| \le C \cdot 2^k|\beta|,$$

for positive j and, for non-positive j, we use

$$|\widehat{K}_t(\beta)| \le \frac{C}{2^k|\beta|};$$

in both cases we get the estimate in (2.21). □

At this point it is convenient to introduce the norm

$$\|\cdot\| = \|\|\cdot\|_{v(2)}\|_{\ell^2}.$$

Since $\mathcal{S}$ commutes with the ℓ^2 norm, we can rewrite (2.19) as

$$\mathcal{S}(\|(K_t\phi \mid 2^{k-1} \le t < 2^k)\|) \le C \cdot \|\phi\|_{\ell^2}.$$

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

By the triangle inequality, we have

$$\mathcal{S}(\|(K_t\phi \mid 2^{k-1} \le t < 2^k)\|) \le \sum_j \mathcal{S}(\|(K_{j,t}\phi \mid 2^{k-1} \le t < 2^k)\|),$$

and we see that our proof will be finished if we prove

$$\mathcal{S}(\|(K_{j,t}\phi \mid 2^{k-1} \le t < 2^k)\|) \le \frac{C}{2^{|j|/2}} \cdot \|\phi\|_{\ell^2}. \tag{2.23}$$

One more piece of notation and we can start. Let

$$I_{j,k} = \{\beta \mid 2^{-j-k-1} \le |\beta| \le 2^{-j-k}\},$$
$$N = \min\{2^{k-1}, 2^{|j|}\}.$$

Choose N equidistributed points between 2^{k-1} and 2^k, say,

$$2^{k-1} = a_1 < a_2 < \cdots < a_N < 2^k.$$

Define the operators A_t by

$$A_t = K_{j,a_{m-1}}; \quad a_{m-1} \le t < a_m.$$

Estimate

$$\begin{aligned}\mathcal{S}(\|(K_{j,t}\phi \mid 2^{k-1} \le t < 2^k)\|) \quad &\le \quad \mathcal{S}(\|(A_t\phi \mid 2^{k-1} \le t < 2^k)\|) \\ &\quad +\mathcal{S}(\|(A_t\phi - K_{j,t}\phi \mid 2^{k-1} \le t < 2^k)\|). \end{aligned} \tag{2.24}$$

Let us bound the first term above. For fixed k, we have, using (2.6) and the fact that S commutes with the ℓ^2 norm,

$$\begin{aligned}\|(A_t\phi \mid 2^{k-1} \le t < 2^k)\| \quad &= \quad \|(K_{j,a_m}\phi \mid m \le N)\| \\ &\le \quad 2 \cdot \mathcal{S}(\|K_{j,a_m}\phi\|_{\ell^2}).\end{aligned}$$

By Parseval's formula, we have

$$\sum_{m \le N} \|K_{j,a_m}\phi\|_{\ell^2}^2 \quad = \quad \sum_{m \le N} \int_{[-1/2,1/2)} |\widehat{K}_{j,a_m}(\beta)\widehat{\phi}(\beta)|^2 \, d\beta$$

by (2.21) and noting that $\widehat{K}_{j,a_m}(\beta)$ is non-zero only if $\beta \in I_{j,k}$,

$$\begin{aligned} &\le \quad C \cdot \sum_{m \le N} \frac{1}{2^{2|j|}} \int_{I_{j,k}} |\widehat{\phi}(\beta)|^2 \, d\beta \\ &\le \quad \frac{C}{2^{|j|}} \int_{I_{j,k}} |\widehat{\phi}(\beta)|^2 \, d\beta, \end{aligned}$$

where we used, in the last step, that $N \le 2^{|j|}$. Summing the above estimate in k we obtain

$$\begin{aligned}\sum_k \|(A_t\phi \mid 2^{k-1} \le t < 2^k)\|^2 \quad &\le \quad \frac{C}{2^{|j|}} \cdot \sum_k \int_{I_{j,k}} |\widehat{\phi}(\beta)|^2 \, d\beta \\ &= \quad \frac{C}{2^{|j|}} \cdot \int_{[-1/2,1/2)} |\widehat{\phi}(\beta)|^2 \, d\beta,\end{aligned}$$

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

which, by Parseval's formula, implies the bound

$$\mathcal{S}(\|(A_t\phi \mid 2^{k-1} \le t < 2^k)\|) \le \frac{C}{2^{|j|/2}} \cdot \|\phi\|_{\ell^2}. \tag{2.25}$$

Let us bound the second term in (2.24). Since $A_t = K_{j,t}$ if $2^{k-1} \le t < 2^k$ and $k \le |j|$, we assume that $k > |j|$, and hence $N = 2^{|j|}$. By (2.8) we can estimate as

$$\begin{aligned} \|(A_t\phi - K_{j,t}\phi \mid 2^{k-1} \le t < 2^k)\|_{v(2)} &\le 2 \cdot \mathcal{S}(\|(A_t\phi - K_{j,t}\phi \mid a_{m-1} \le t < a_m)\|_{v(2)}) \\ &= 2 \cdot \mathcal{S}(\|(K_{j,t}\phi \mid a_{m-1} \le t < a_m)\|_{v(2)}). \end{aligned}$$

For fixed m, we have

$$\begin{aligned} &\|(K_{j,t}\phi \mid a_{m-1} \le t < a_m)\|_{v(2)} \\ &\le \|(K_{j,t}\phi \mid a_{m-1} \le t < a_m)\|_{v(1)} \\ &\le \sum_{a_{m-1} \le t < a_m} |K_{j,t}\phi - K_{j,t+1}\phi| \\ &\le (a_m - a_{m-1})^{1/2} \cdot \left(\sum_{a_{m-1} \le t < a_m} (K_{j,t}\phi - K_{j,t+1}\phi)^2 \right)^{1/2}, \end{aligned}$$

where in the last step we used Cauchy's inequality. If we put the above two estimates together, take the ℓ^2 norm and use that $\mathcal{S}$ commutes with the ℓ^2 norm, we get

$$\begin{aligned} &\|(A_t\phi - K_{j,t}\phi \mid 2^{k-1} \le t < 2^k)\| \\ &\le 2 \cdot \mathcal{S}\left((a_m - a_{m-1})^{1/2} \cdot \left(\sum_{a_{m-1} \le t < a_m} \|K_{j,t}\phi - K_{j,t+1}\phi\|_{\ell^2}^2 \right)^{1/2} \right). \end{aligned}$$

By Parseval's formula, and recalling that $a_m - a_{m-1} = 2^k/N$, we have

$$\begin{aligned} &\sum_{m \le N} (a_m - a_{m-1}) \cdot \sum_{a_{m-1} \le t < a_m} \|K_{j,t}f - K_{j,t+1}f\|_{\ell^2}^2 \\ &= \sum_{m \le N} \frac{2^k}{N} \cdot \sum_{a_{m-1} \le t < a_m} \int_{[-1/2,1/2)} |\widehat{K}_{j,t}(\beta) - \widehat{K}_{j,t+1}(\beta)|^2 |\widehat{f}|^2 \, d\beta \end{aligned}$$

by (2.22) and noting that $\widehat{K}_{j,t}(\beta)$ is non-zero only if $\beta \in I_{j,k}$,

$$\begin{aligned} &\le C \cdot \sum_{m \le N} \frac{2^k}{N} \cdot \frac{2^k}{N} \cdot \frac{1}{2^{2k}} \cdot \int_{I_{j,k}} |\widehat{\phi}|^2 \, d\beta \\ &= \frac{C}{N} \cdot \int_{I_{j,k}} |\widehat{\phi}|^2 \, d\beta. \end{aligned}$$

Summing the above estimate in k we obtain, as before,

$$\begin{aligned} \sum_k \|(A_t\phi - K_{j,t}\phi \mid 2^{k-1} \le t < 2^k)\|^2 &\le \frac{C}{N} \cdot \sum_k \int_{I_{j,k}} |\widehat{\phi}(\beta)|^2 \, d\beta \\ &= \frac{C}{2^{|j|}} \cdot \int_{[-1/2,1/2)} |\widehat{\phi}(\beta)|^2 \, d\beta, \end{aligned}$$

which, by Parseval's formula, implies the bound

$$\mathcal{S}(\|(A_t\phi - K_{j,t}\phi \mid 2^{k-1} \le t < 2^k)\|) \le \frac{C}{2^{|j|/2}} \cdot \|\phi\|_{\ell^2}. \tag{2.26}$$

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

Now, because of (2.24), the estimates in (2.25) and (2.26) prove (2.23), and so we are done. □

Let us note that an important consequence of the proof is a generalization of Theorem 2.4. Denote by E_{2^k} what we called E_k before, and define the operators E_t by

$$E_t = E_{2^k}; \quad 2^{k-1} \le t < 2^k.$$

THEOREM 2.17.

$$\| \|(M_t\phi - E_t\phi)\|_{v(2)} \|_{\ell^2} \le C \cdot \|\phi\|_{\ell^2}.$$

3. *Weak* (1, 1) *results*

We begin this section with a general result that will be used to establish the weak (1, 1) inequalities for a variety of operators considered in the earlier sections. While the argument uses the standard Calderón–Zygmund decomposition, it uses an ℓ^2 estimate (or, more generally, an ℓ^ϱ estimate for some $\varrho \ge 1$) for the 'bad' function, and hence is different from the standard argument. Note that for square functions, ℓ^2 information is often more natural.

THEOREM 3.1. *Let A be a sublinear operator. Assume that A satisfies:*

(i) *there is a constant c so that for any $\phi : \mathbb{Z}^d \to \mathbb{R}$ of finite support*

$$|\{|A\phi| > 1\}| \le c\|\phi\|_{\ell^2}^2;$$

(ii) *there is a constant C and an $\varrho \ge 1$ such that if (B_j) is a collection of pairwise disjoint dyadic cubes in $\mathbb{Z}^d$, and $b : \mathbb{Z}^d \to \mathbb{R}$ is a function of finite support with $b = \sum_j b_j$, such that each b_j is supported on B_j, and $\sum_{x \in B_j} b_j(x) = 0$ for each j, then we have a 'reverse ℓ^ϱ inequality'; that is*

$$\left| A\left(\sum_j b_j(x) \right) \right| \le C \left(\sum_{j=1}^{\infty} |A(b_j)(x)|^\varrho \right)^{1/\varrho},$$

for all $x \notin \cup_j \tilde{B}_j$, where for any cube, B, $\tilde{B}$ denotes the cube with the same center as B, but with each dimension expanded three times;

(iii) *there is a constant C so that if B is a dyadic cube and the function $b : \mathbb{Z}^d \to \mathbb{R}$ is supported on B, $\sum_{x \in \mathbb{Z}^d} b(x) = 0$ and $\|b\|_{\ell^1} \le 4^d \cdot |B|$ then we have*

$$\sum_{x \notin \tilde{B}} |Ab(x)|^\varrho \le C|B|.$$

If all these properties are satisfied then there is a constant c so that for any $\phi : \mathbb{Z}^d \to \mathbb{R}$ of finite support

$$|\{|A\phi| > 1\}| \le c\|\phi\|_{\ell^1}.$$

Further, the same result holds if we replace $\mathbb{Z}^d$ by $[0, 1)^d$ or by $\mathbb{R}^d$ and use L^p norms in place of ℓ^p norms.

Remark. When $\varrho = 1$, this theorem is essentially the Benedek–Calderón–Panzone principle. In this case, since A is sublinear, the condition in (ii) is automatically satisfied.

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

Proof. We will give the details only in the case of $\mathbb{Z}^d$. The proof in the other cases is exactly the same. We begin with the standard Calderón–Zygmund decomposition. We state the following lemma without proof (see [**26**]).

LEMMA 3.2. (Calderón–Zygmund decomposition.) *Any $\phi : \mathbb{Z}^d \to \mathbb{R}$ of finite support can be written in the form $\phi = g + b$ where:*

1. $\|g\|_{\ell^2}^2 \le \|\phi\|_{\ell^1}$.
2. $b = \sum_j b_j$ *where each b_j satisfies:*
   (a) *b_j is supported in a dyadic cube B_j;*
   (b) $\sum_{x\in\mathbb{Z}^d} b_j(x) = 0$ *for each i;*
   (c) $\|b_j\|_{\ell^1} \le 4^d |B_j|$;
   (d) $\sum_j |B_j| \le \|\phi\|_{\ell^1}$.

Now let $\phi = g + b$ as given above. Then we have

$$|\{|A\phi| > 2\}| \le |\{|Ag| > 1\}| + |\{|Ab| > 1\}|.$$

For the first piece we use (i) and 1. above to obtain

$$|\{|Ag| > 1\}| \le c\|g\|_{\ell^2}^2 \le c\|\phi\|_{\ell^1}.$$

For the second piece we break the set into two pieces. We can estimate the second piece by

$$|\{|Ab| > 1\}| \le |\{x : |Ab| > 1, x \in \cup_j \tilde{B}_j\}| + |\{x : |Ab| > 1, x \notin \cup_j \tilde{B}_j\}|.$$

The first of these is easy; we have, by 2(d),

$$\begin{aligned} |\{x : |Ab| > 1, x \in \cup_j \tilde{B}_j\}| &\le |\cup_j \tilde{B}_j| \\ &\le \sum_{j=1}^{\infty} |\tilde{B}_j| \\ &\le 3^d \sum_{j=1}^{\infty} |B_j| \\ &\le 3^d \sum_{j=1}^{\infty} \sum_{x\in B_j} |\phi(x)| \\ &\le 3^d \|\phi\|_{\ell^1}. \end{aligned}$$

For the second piece we need to use the other properties of A in the statement of the theorem. We have

$$|\{x : |Ab| > 1, x \notin \cup_j \tilde{B}_j\}| \le \sum_{x\notin\cup_j \tilde{B}_j} |Ab(x)|^{\varrho}$$

by (ii)

$$\begin{aligned} &\le c \sum_{x\notin\cup_j \tilde{B}_j} \sum_{j=1}^{\infty} |Ab_j(x)|^{\varrho} \\ &\le c \sum_{j=1}^{\infty} \sum_{x\notin\tilde{B}_j} |Ab_j(x)|^{\varrho} \end{aligned}$$

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

by (iii)

$$\leq \quad c \sum_{j=1}^{\infty} C \cdot |B_j|$$

by 2(d)

$$\leq \quad c\|\phi\|_{\ell^1}. \qquad \square$$

We now want to apply this result to the operators considered earlier. In each case we have already proven that the operator is strong type $(2, 2)$. Hence we already know that condition (i) is satisfied.

First we prove a result which will imply that the square function considered in Theorem 2.1 is weak type $(1, 1)$. In fact, our formulation will be more general, but we do not achieve yet the generality of Theorem 2.2; we restrict the definition of the operator D_n of Theorem 2.2 by assuming that for each n, the functions $a_n(x)$ and $b_n(x)$ are constants. As we will see, in this more restrictive setting, we can use Theorem 3.1 with $\varrho = 1$. Later, when we prove the weak $(1, 1)$ analog of Theorem 2.2 (Theorem 3.5), we will need to use the full strength of Theorem 3.1; we will take $\varrho = 2$.

Let (R_n) be a sequence of intervals in $[0, 1)$ all containing 0 and $|R_n| = 2^{-n}$. Let $D_n f(x) = 2^n \int_{R_n} f(x + t)\, dt$, and let $f_n = \mathbb{E}_n f$ be the standard dyadic martingale on $[0, 1)$.

THEOREM 3.3. *The square function*

$$S(f)(x) = \mathcal{S}(D_n f(x) - f_n(x)),$$

where D_n and f_n are defined above, is a weak type $(1, 1)$ operator.

Proof. We will use Theorem 3.1 with $A = S$ and $\varrho = 1$.

Condition (i) holds because of Theorem 2.2. We do not have to check (ii) since S is subadditive. To check (iii), let B be a dyadic interval of length $2^{-\ell}$, and let b be a function supported on B, of mean value 0 and $\|b\|_1 \leq 4|B|$. Since, for every n, $b_n(x) = 0$ if $x \in \tilde{B}^c$, we have

$$Sb(x) = \sqrt{\sum_n |D_n b(x)|^2}$$

for $x \in \tilde{B}^c$. Since $Sb(x) \leq \sum_n |D_n b(x)|$, we just need to show

$$\int_{\tilde{B}^c} \sum_n |D_n b(x)|\, dx = \sum_n \int_{\tilde{B}^c} |D_n b(x)|\, dx \leq C\|b\|_{L^1}.$$

Let $x \in \tilde{B}^c$. Since b is supported on B we have $D_n b(x) = 0$ if $n \geq \ell$. So we just need to consider terms with $n < \ell$. Since the mean value of b is 0, we have $D_n b(x) = 0$ if B is contained in the interval $x + R_n$. This implies that, with $B = [c, d)$, for $n < \ell$, $D_n b(x)$ may be non-zero only on one of the intervals $[c, d) - 2^n$ or $[c, d) + 2^n$. As a consequence, since $|D_n b(x)| \leq 2^n \cdot \|b\|_{L^1}$, we have

$$\int_{\tilde{B}^c} |D_n b(x)|\, dx \leq 2 \cdot 2^n \|b\|_{L^1} \cdot 2^{-\ell}.$$

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

Summing this over n, $n < \ell$, we obtain

$$\int_{\tilde{B}^c} \sum_n |D_n b(x)|\, dx \le C\|b\|_{L^1} \le C \cdot |B|,$$

which finishes our proof. □

The following theorem is proved in much the same way, but we need to use Theorem 2.4 to check condition (i) in Theorem 3.1.

THEOREM 3.4. *The square function considered in Theorem 2.4 and defined by*

$$S\phi(r) = \left(\sum_{n=1}^{\infty} |M_{2^n}\phi(r) - E_n\phi(r)|^2 \right)^{1/2}$$

is a weak type $(1, 1)$ *operator.*

Let us now turn to applications of Theorem 3.1 where we need to take $\varrho > 1$.

Let $D_n f(x) = 2^n \int_{R_n(x)} f(y)\, dy$, where for each positive integer n, $R_n(x)$ is an interval of length 2^{-n} containing x, chosen in a measurable fashion. Let $f_n = \mathbb{E}_n f$ be the standard dyadic martingale on $[0, 1)$.

THEOREM 3.5. *The square function*

$$S(f)(x) = \mathcal{S}(D_n f(x) - f_n(x)),$$

where D_n *and* f_n *are defined above, is a weak type* $(1, 1)$ *operator.*

Proof. We will use Theorem 3.1 with $A = S$ and $\varrho = 2$.

Condition (i) holds because of Theorem 2.2.

To check (ii), let B_j, b_j and $b = \sum b_j$ be as stipulated by this property, and fix $x \in (\cup \tilde{B}_j)^c$. Observe that, because of the position of x and the assumptions on b_j and B_j, we have $(b_j)_n(x) = 0$ for each j, and hence $b_n(x) = 0$. It follows that $Sb(x) = \sqrt{\sum (D_n b(x))^2}$. Notice that for each n, there are at most two j's for which $D_n b_j(x) \neq 0$. Indeed, $D_n b_j(x) \neq 0$ could happen only if B_j contains the starting point or endpoint of $R_n(x)$. It follows that $(D_n b(x))^2 \le 4 \sum_j (D_n b_j(x))^2$. Summing this inequality in n, we obtain

$$\begin{aligned} (Sb(x))^2 &= \sum_n (D_n b(x))^2 \\ &\le 4 \sum_j \sum_n (D_n b_j(x))^2 \\ &= 4 \sum_j (Sb_j(x))^2, \end{aligned}$$

which is equivalent to what we wanted to prove.

We now check property (iii). Let B be a dyadic interval, let b be supported on B with mean value 0 and $\|b\|_1 \le 4|B|$, and let $x \in \tilde{B}^c$. We denote by $\delta = \delta(x)$ the distance between x and B. As before, we observe that $Sb(x) = \sqrt{\sum D_n^2 b(x)}$. Since b is supported

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

on B, we have $D_n b(x) = 0$ if $2^{-n} \le \delta$. It follows that $S^2 b(x) = \sum_{2^{-n}>\delta} D_n^2 b(x)$. Since $D_n b(x) \le 2^n \|b\|_{L^1}$, we can estimate

$$S^2 b(x) \le \sum_{2^{-n}>\delta} 2^{2n} \|b\|_{L^1}^2 \le C \cdot \|b\|_{L^1}^2 \cdot \frac{1}{\delta^2}.$$

But then

$$\begin{aligned} \sum_{x\in\tilde{B}^c} S^2 b(x) &\le \sum_{\{x|\delta(x)\ge|B|\}} C \cdot \|b\|_{L^1}^2 \cdot \frac{1}{\delta^2(x)} \\ &\le C \cdot \|b\|_{L^1}^2 \sum_{y\ge|B|} y^{-2} \\ &\le C \cdot \|b\|_{L^1}^2 \cdot |B|^{-1} \\ &\le C \cdot |B|, \end{aligned}$$

which was to be proven. □

THEOREM 3.6. *Let $R_1 \subset R_2 \subset \cdots$ denote a nested sequence of intervals in $\mathbb{Z}$, such that $|R_t| = t$ and $R_1 = \{0\}$. Let $M_t\phi(r) = (1/t)\sum_{j\in R_t} \phi(r+j)$.*

(i) *If $2 < \varrho < \infty$ then the operator*

$$V_\varrho \phi(r) = \|(M_t\phi)(r)\|_{v(\varrho)} \tag{3.1}$$

maps ℓ^1 into weak ℓ^1.

(ii) *For any sequence $n_1 < n_2 < \cdots$ the operator*

$$O\phi(r) = \|(M_t\phi)(r)\|_{o(2)} \tag{3.2}$$

maps ℓ^1 into weak ℓ^1.

(iii) *Further, if the averages M_t denote the ergodic averages*

$$M_t f(x) = \frac{1}{t}\sum_{k\in R_t} f(\tau^k x),$$

then both operators map $L^1(X)$ to weak $L^1(X)$.

Proof. We prove (i) only. The proof of (ii) is similar, and (iii) follows by the transference principle.

In Theorem 2.13 we have already shown that condition (i) of Theorem 3.1 holds, and that (ii) of Theorem 3.1 holds is seen in much the same way as in the proof of Theorem 3.5.

It now remains to prove that (iii) of Theorem 3.5 holds. Let B be an interval, let b be supported on B with mean value 0 and $\|b\|_{\ell^1} \le 4|B|$, and let $r \in \tilde{B}^c$. We denote by $\delta = \delta(r)$ the distance between r and B. Let us choose $n_1 < n_2 < \cdots$ so that

$$V_\varrho b(r) = \left(\sum_{k=2}^{\infty} |M_{n_k} b(r) - M_{n_{k-1}} b(r)|^\varrho\right)^{1/\varrho}.$$

Since $M_n b(r) = 0$ if $n \le \delta$, we can assume that either $n_1 > \delta$ or $n_1 \le \delta < n_2$. We just deal with the second case; the first case is a bit simpler and handled similarly.

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

Since in this case $M_{n_1}b(r) = 0$, we can estimate

$$\begin{aligned}
&V_\varrho b(r)\\
&= \left(|M_{n_2}b(r)|^\varrho + \sum_{k=3}^{\infty} |M_{n_k}b(r) - M_{n_{k-1}}b(r)|^\varrho\right)^{1/\varrho}\\
&\le |M_{n_2}b(r)| + \sum_{k=3}^{\infty} |M_{n_k}b(r) - M_{n_{k-1}}b(r)|\\
&\le |M_{n_2}b(r)| + \sum_{k=3}^{\infty}\left(\left|\left(\frac{1}{n_k} - \frac{1}{n_{k-1}}\right)\sum_{j\in R_{n_{k-1}}} b(r+j)\right| + \left|\frac{1}{n_k}\sum_{j\in R_{n_k}\setminus R_{n_{k-1}}} b(r+j)\right|\right)\\
&\le \frac{1}{n_2}\|b\|_{\ell^1} + \sum_{k=3}^{\infty}\left(\frac{1}{n_k} - \frac{1}{n_{k-1}}\right)\|b\|_{\ell^1} + \frac{1}{n_2}\sum_{k=3}^{\infty}\sum_{j\in R_{n_k}\setminus R_{n_{k-1}}} |b(r+j)|\\
&\le \frac{c}{n_2}\|b\|_{\ell^1}\\
&\le \frac{c}{\delta}\|b\|_{\ell^1}.
\end{aligned}$$

We now have

$$\begin{aligned}
\sum_{r\in\tilde{B}^c} V_\varrho^\varrho b(r) &\le \sum_{\{r|\delta(r)\ge|B|\}} C\cdot\|b\|_{\ell^1}^\varrho\cdot\frac{1}{\delta^\varrho(r)}\\
&\le C\cdot\|b\|_{\ell^1}^\varrho\sum_{y\ge|B|} y^{-\varrho}\\
&\le C\cdot\|b\|_{\ell^1}^\varrho\cdot|B|^{1-\varrho}\\
&\le C\cdot|B|,
\end{aligned}$$

which completes our proof. □

We finally mention the following result, the proof of which is similar to the above.

THEOREM 3.7. *Let E_t be the operator as defined in Theorem 2.17. Then the operator $\|(M_t\phi - E_t\phi)\|_{v(2)}$ is weak type* (1, 1).

4. *BMO and L^p results*

A general principle. In trying to prove that the square function operator

$$Sf(x) = \sqrt{\sum_n (M_n f(x) - M_{n-1}f(x))^2}$$

maps $\ell^\infty \to BMO$, we notice that $|Sf(x) - Sf(y)| \le Bf(x, y)$, where

$$Bf(x, y) = \sqrt{\sum_n (M_n f(x) - M_{n-1}f(x) - (M_n f(y) - M_{n-1}f(y)))^2}.$$

If $f = g + h$, then this operator B certainly satisfies

$$Bf(x, y) \le Bg(x, y) + Bh(x, y)$$

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

and

$$Bf(x, y) \leq 2 \cdot \max\{Sf(x), Sf(y)\}.$$

For an interval $I \subset \mathbb{Z}$, we denote by $\tilde{I}$ the interval with the same center as I but of length $3|I|$.

Similar to the previous section, and motivated by the above observations, we begin with a general result that will imply that many of the operators we considered earlier map ℓ^∞ (or L^∞) to BMO or dyadic-BMO.

THEOREM 4.1. *Let A be an operator defined on ℓ^∞. Assume the following five properties:*

(i) *the operator A is a bounded $\ell^2 \to \ell^2$ operator;*

(ii) *there is an operator B which satisfies $|Af(x) - Af(y)| \leq C \cdot Bf(x, y)$ for some constant C and every f, x, y, and which satisfies the following three properties;*

(iii) *there is a constant C such that if $f = g + h$, then*

$$Bf(x, y) \leq C \cdot (Bg(x, y) + Bh(x, y));$$

(iv) *there is a constant C such that we have*

$$Bf(x, y) \leq C \cdot \max\{Af(x), Af(y)\};$$

(v) *there is a constant C so that for any interval $I \subset \mathbb{Z}$, and $h \in \ell^\infty$ supported on $\mathbb{Z} \setminus \tilde{I}$, we have*

$$\sup_{x,y \in I} Bh(x, y) \leq C \cdot \|h\|_{\ell^\infty}.$$

Then there is a constant C such that $\|Af\|_{BMO} \leq C\|f\|_\infty$. The constant C depends only on the constants in (ii)–(v) and the $\ell^2 \to \ell^2$ bound of A.

Remark 4.2. A similar result is true if we restrict the intervals I in (v) to dyadic intervals, but then the conclusion is that A maps ℓ^∞ to dyadic-BMO.

Remark 4.3. Theorem 4.1 and the result in Remark 4.2 extends to the cases when $\mathbb{Z}$ is replaced by $\mathbb{R}$ or $\mathbb{T}$ (the torus).

Proof. We need to show that for any interval $I \subset \mathbb{Z}$ we have

$$\frac{1}{|I|} \sum_{x \in I} \left| Af(x) - \frac{1}{|I|} \sum_{y \in I} Af(y) \right| \leq C\|f\|_\infty.$$

We have

$$\begin{aligned} \frac{1}{|I|} \sum_{x \in I} \left| Af(x) - \frac{1}{|I|} \sum_{y \in I} Af(y) \right| &= \frac{1}{|I|} \sum_{x \in I} \left| \frac{1}{|I|} \sum_{y \in I} Af(x) - Af(y) \right| \\ &\leq \frac{1}{|I|^2} \sum_{x \in I} \sum_{y \in I} |Af(x) - Af(y)|. \end{aligned}$$

By (ii), we need to show, for any interval $I \subset \mathbb{Z}$,

$$\frac{1}{|I|^2} \sum_{x \in I} \sum_{y \in I} Bf(x, y) \leq C\|f\|_\infty. \tag{4.1}$$

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

Fix an interval I. Let $g(x) = f(x)\chi_{\tilde{I}}(x)$, the restriction of f to $\tilde{I}$, and let $h(x) = f(x) - g(x)$, the restriction of f to $\mathbb{Z} \setminus \tilde{I}$. By the subadditive property in (iii), we have

$$Bf(x, y) \leq C \cdot (Bg(x, y) + Bh(x, y)).$$

It follows that it is enough to prove the following two inequalities:

$$\frac{1}{|I|^2} \sum_{x \in I} \sum_{y \in I} Bg(x, y) \leq C\|g\|_\infty, \tag{4.2}$$

with a constant C depending on the ℓ^2 bound of the operator A; and

$$\sup_{x,y \in I} Bh(x, y) \leq C\|h\|_\infty. \tag{4.3}$$

Only (4.2) requires proof since (4.3) is just assumption (v) in our theorem.

Since, by assumption (iv), $Bg(x, y) \leq C \cdot \max\{Ag(x), Ag(y)\}$, we need to prove

$$\frac{1}{|I|} \sum_{x \in I} Ag(x) \leq C\|g\|_\infty. \tag{4.4}$$

Since, on a probability space, $\|\cdot\|_{L^1} \leq \|\cdot\|_{L^2}$, we have

$$\frac{1}{|I|} \sum_{x \in I} Ag(x) \leq \sqrt{\frac{1}{|I|} \sum_{x \in I} |Ag(x)|^2}.$$

Since, by the assumption in (i) A is a type $(2, 2)$ operator, we have

$$\sum_{x \in I} |Ag(x)|^2 \leq \|Ag\|_{\ell^2}^2 \leq C\|g\|_{\ell^2}^2.$$

The inequality in (4.4) follows, since g is supported on $\tilde{I}$, an interval three-times the length of I, and $\|g\|_{\ell^2}^2 \leq |\tilde{I}|\|g\|_\infty^2$, so

$$\sqrt{\frac{1}{|I|} \sum_{x \in I} |Ag(x)|^2} \leq C\|g\|_\infty. \qquad \square$$

Applications of Theorem 4.1. Let us introduce some notation which we will use for the rest of this section.

Let R_n denote a nested sequence of intervals in $\mathbb{Z}$, with $|R_n| = n$, $R_1 = \{0\}$, and define $M_n f(x) = (1/n) \sum_{j \in R_n} f(x + j)$. Fix $\varrho > 2$ and a sequence $n_1 < n_2, \ldots$. Let us define the operators A_1, A_2 and A_3 by

$$\begin{aligned}
A_1 f(x) &= \mathcal{S}(M_{n_k} f(x) - M_{n_{k-1}} f(x)) = \left(\sum_{k=1}^{\infty} |M_{n_k} f(x) - M_{n_{k-1}} f(x)|^2 \right)^{1/2}; \\
A_2 f(x) &= \mathcal{O}(M_n f(x)) = \left(\sum_{k=1}^{\infty} \sup_{n_{k-1} \leq n \leq m < n_k} |M_n f(x) - M_m f(x)|^2 \right)^{1/2}; \\
A_3 f(x) &= \mathcal{V}_\varrho(M_n f(x)) = \sup_{(n(i))} \left(\sum_i |M_{n(i)} f(x) - M_{n(i+1)} f(x)|^\varrho \right)^{1/\varrho}.
\end{aligned} \tag{4.5}$$

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

Finally, for $f \in L^1([0,1))$, define the operator A_4 by

$$A_4 f(x) = \mathcal{S}(D_n f(x) - f_n(x)) = \left(\sum_{n=1}^{\infty} |D_n f(x) - f_n(x)|^2 \right)^{1/2}, \tag{4.6}$$

where D_n and f_n are as in (1.7).

THEOREM 4.4.

(a) *The operators A_1, A_2 and A_3 defined in (4.5) are bounded $L^\infty \to BMO$ operators. In the cases A_1 and A_2, the $L^\infty \to BMO$ bounds depend only on the $\ell^2 \to \ell^2$ bounds, but do not depend on the sequence (n_k). In case of A_3, the $L^\infty \to BMO$ bound depends only on the $\ell^2 \to \ell^2$ bound and $\varrho > 2$.*

(b) *The operator A_4 defined in (4.6) is a bounded $L^\infty \to$ dyadic-BMO operator. The $L^\infty \to$ dyadic-BMO norm of A_4 depends only on its $\ell^2 \to \ell^2$ bound.*

Proof of part (a) of Theorem 4.4. We just give the proof for the operator A_3, but formulate the proof so that it will apply equally well to A_2 (just change ϱ in the proof to 2). With only trivial changes, the result for A_1 follows in the same way.

We want to apply Theorem 4.1 with $A = A_3$ and

$$Bf(x,y) = \|M_n f(x) - M_n f(y)\|_{v(\varrho)}.$$

By Theorem 2.13, the property in (i) is satisfied. Properties (ii)–(iv) follow easily from the triangle inequality for the seminorm $\|\cdot\|_{v(\varrho)}$. It remains to check the condition in (v).

Let $I \subset \mathbb{Z}$ and h be supported on $\mathbb{Z} \setminus \tilde{I}$; we can assume that $\|h\|_{\ell^\infty} \le 1$. Let $x, y \in I$. By the definition of the norm $\|\cdot\| = \|\cdot\|_{v(\varrho)}$, there are interlacing sequences $(m(k))$ and $(n(k))$ (interlacing means that $m(1) < n(1) \le m(2) < n(2) \le \cdots$), so that

$$\begin{aligned} &\|M_n h(x) - M_n h(y)\| \\ &\quad \le \ 2 \cdot \left(\sum_k |(M_{n(k)}h(x) - M_{m(k)}h(x)) - (M_{n(k)}h(y) - M_{m(k)}h(y))|^\varrho \right)^{1/\varrho}. \end{aligned}$$

Since the function h is supported outside the interval $\tilde{I}$, we have $M_t h(z) = 0$ for $t \le |I|$ and $z \in I$. Therefore, we can assume $m(1) > |I|$. Let us write

$$\begin{aligned} &\sum_k |(M_{n(k)}h(x) - M_{m(k)}h(x)) - (M_{n(k)}h(y) - M_{m(k)}h(y))|^\varrho \\ &\quad = \sum_{n(k)-m(k)>|I|} |(M_{n(k)}h(x) - M_{m(k)}h(x)) - (M_{n(k)}h(y) - M_{m(k)}h(y))|^\varrho \\ &\qquad + \sum_{n(k)-m(k)\le|I|} |(M_{n(k)}h(x) - M_{m(k)}h(x)) - (M_{n(k)}h(y) - M_{m(k)}h(y))|^\varrho. \end{aligned}$$

By the triangle inequality for the ℓ^ϱ-norm, we need to establish the following two inequalities:

$$\left(\sum_{n(k)-m(k)>|I|} |(M_{n(k)}h(x) - M_{m(k)}h(x)) - (M_{n(k)}h(y) - M_{m(k)}h(y))|^\varrho \right)^{1/\varrho} \le C \tag{4.7}$$

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

and

$$\left(\sum_{n(k)-m(k)\le|I|}|(M_{n(k)}h(x)-M_{m(k)}h(x))-(M_{n(k)}h(y)-M_{m(k)}h(y))|^{\varrho}\right)^{1/\varrho}\le C. \quad (4.8)$$

Let us prove (4.7). We have, by the triangle inequality,

$$\begin{aligned}\left(\sum_{n(k)-m(k)>|I|}|(M_{n(k)}h(x)-M_{m(k)}h(x))-(M_{n(k)}h(y)-M_{m(k)}h(y))|^{\varrho}\right)^{1/\varrho}\\ \le\left(\sum_{n(k)-m(k)>|I|}|M_{n(k)}h(x)-M_{n(k)}h(y)|^{\varrho}\right)^{1/\varrho}\\ +\left(\sum_{n(k)-m(k)>|I|}|M_{m(k)}h(x)-M_{m(k)}h(y)|^{\varrho}\right)^{1/\varrho}.\end{aligned}$$

We just need to prove

$$\left(\sum_{n(k)-m(k)>|I|}|M_{n(k)}h(x)-M_{n(k)}h(y)|^{\varrho}\right)^{1/\varrho}\le C \quad (4.9)$$

with a constant C independent of the increasing sequence $(n(k))$. Observe that, for every n,

$$|M_nh(x)-M_nh(y)|\le 2\cdot\frac{|I|}{n}. \quad (4.10)$$

Indeed,

$$\begin{aligned}|M_nh(x)-M_nh(y)| &= \left|\frac{1}{n}\sum_{j\in(x+R_n)\setminus(y+R_n)}h(j)-\frac{1}{n}\sum_{j\in(y+R_n)\setminus(x+R_n)}h(j)\right|\\ &\le \frac{1}{n}\sum_{j\in D}|h(j)|,\end{aligned}$$

where D is the symmetric difference of the sets $x+R_n$ and $y+R_n$. The estimate in (4.10) follows by noting that $|D|\le 2|x-y|\le 2|I|$.

Since the sequences $(n(k))$ and $(m(k))$ are interlacing and because of the assumption $n(k)-m(k)>|I|$, we have that $n(k)\ge k|I|$. But then

$$\sum_{n(k)-m(k)>|I|}|M_{n(k)}h(x)-M_{n(k)}h(y)|^{\varrho}\le 2\sum_k\left(\frac{|I|}{k\cdot|I|}\right)^{\varrho}<\infty,$$

since $\varrho>1$. This proves (4.9), and hence (4.7).

Let us now prove (4.8). This time, we estimate, using the triangle inequality,

$$\begin{aligned}\left(\sum_{n(k)-m(k)\le|I|}|(M_{n(k)}h(x)-M_{m(k)}h(x))-(M_{n(k)}h(y)-M_{m(k)}h(y))|^{\varrho}\right)^{1/\varrho}\\ \le\left(\sum_{n(k)-m(k)\le|I|}|M_{n(k)}h(x)-M_{m(k)}h(x)|^{\varrho}\right)^{1/\varrho}\\ +\left(\sum_{n(k)-m(k)\le|I|}|M_{n(k)}h(y)-M_{m(k)}h(y)|^{\varrho}\right)^{1/\varrho}.\end{aligned}$$

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

We just need to prove

$$\left(\sum_{n(k)-m(k)\le |I|} |M_{n(k)}h(x) - M_{m(k)}h(x)|^{\varrho} \right)^{1/\varrho} \le C \tag{4.11}$$

with a constant C depending on ϱ only.

Observe that, for every n, m, $n > m$,

$$|M_n h(x) - M_m h(x)| \le 2 \cdot (n-m)\frac{1}{n}. \tag{4.12}$$

Indeed,

$$\begin{aligned} |M_n h(x) - M_m h(x)| &= \left| \left(\frac{1}{n} - \frac{1}{m}\right) \sum_{j\in (x+R_n)\cap (x+R_m)} h(j) + \frac{1}{n} \sum_{j\in (x+R_n)\setminus (x+R_m)} h(j) \right| \\ &\le \left(\frac{1}{m} - \frac{1}{n}\right) \sum_{j\in x+R_m} |h(j)| + \frac{1}{n} \sum_{j\in (x+R_n)\setminus (x+R_m)} |h(j)|. \end{aligned}$$

The estimate in (4.12) follows by noting that $|(x + R_n) \setminus (x + R_m)| = n - m$.

As a consequence of (4.12) and Hölder's inequality, we have

$$\begin{aligned} |M_{n(k)}h(x) - M_{m(k)}h(x)| &\le 2 \cdot (n(k) - m(k))\frac{1}{n(k)} \le 2 \sum_{m(k)\le q < n(k)} \frac{1}{q} \\ &\le 2(n(k) - m(k))^{1-1/\varrho} \left(\sum_{m(k)\le q<n(k)} \frac{1}{q^{\varrho}} \right)^{1/\varrho} \\ &\le 2|I|^{1-1/\varrho} \left(\sum_{m(k)\le q<n(k)} \frac{1}{q^{\varrho}} \right)^{1/\varrho}. \end{aligned}$$

where, in the last step, we used $n(k) - m(k) \le |I|$. Taking the ϱth power, and summing in k, $n(k) - m(k) \le |I|$, we get

$$\begin{aligned} \sum_{n(k)-m(k)\le |I|} |M_{n(k)}h(x) - M_{m(k)}h(x)|^{\varrho} &\le 2^{\varrho}|I|^{\varrho-1} \sum_{q\ge |I|} \frac{1}{q^{\varrho}} \\ &= C < \infty, \end{aligned}$$

which proves (4.11) and hence (4.8). □

Proof of part (b) of Theorem 4.4. We want to apply Theorem 4.1 (more precisely, its variant for $\mathbb{T}$) with $A = A_4$ and $Bf(x, y) = \|(D_n f - f_n)(x) - (D_n f - f_n)(y)\|_{\ell^2}$. By Theorem 2.1, the property in (i) is satisfied. Properties (ii)–(iv) follow easily from the triangle inequality for the seminorm $\|\cdot\|_{\ell^2}$. It remains to check the condition in (v).

Let $I \subset \mathbb{T} = [0, 1)$ be a dyadic interval, say of length 2^{-k}, and let h be supported on $\mathbb{T} \setminus \tilde{I}$; we can assume that $\|h\|_{L^\infty} \le 1$. Since the function h is supported outside the interval $\tilde{I}$, we have $D_n h(z) = 0$ and $h_n = 0$ for $2^{-n} \le |I| = 2^{-k}$ and $z \in I$. Therefore, we need to prove

$$\sup_{x,y\in I} \left(\sum_{n>k} ((D_n h(x) - h_n(x)) - (D_n h(y) - h_n(y)))^2 \right)^{1/2} \le C,$$

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

with an absolute constant C. Note that $h_n(x) = h_n(y)$ for $n \geq k$. Since the ℓ^1 norm dominates the ℓ^2 norm, it is enough to prove

$$\sup_{x,y\in I} \sum_{n>k} |D_n h(x) - D_n h(y)| \leq C. \tag{4.13}$$

Denoting by Δ_n the symmetric difference of the intervals $[x, x+2^{-n})$ and $[y, y+2^{-n})$, we can estimate as

$$\begin{aligned} |D_n h(x) - D_n h(y)| &\leq 2^n \int_{\Delta_n} |h(z)|\, dz \\ &\leq 2^n \cdot 2|x-y| = 2^{n+1-k}, \end{aligned}$$

where we used that $\|h\|_{L^\infty} \leq 1$. Summing this over $n < k$, we see we can take $C = 2$ in (4.13). □

Remark 4.5. For the BMO result to hold, it is important that the operators D_n in the definition of A_4 are not too general. In particular, the square function considered in Theorem 2.2 may fail to map L^∞ to BMO. We show this with the following example.

Let f_k be the usual dyadic martingale on $[0,1)$ and let $D_k f(x) = 2^k \int_{I_k(x)} f(t)\, dt$, where $I_k(x)$ is an interval of length $1/2^k$ which contains the point x. For certain choices of $(I_k(x))$, the square function

$$Sf(x) = \left(\sum_{k=0}^{\infty} |D_k f(x) - f_k(x)|^2 \right)^{1/2}$$

does not map L^∞ into BMO.

To see this, let P and N denote two disjoint subsets of $[\frac{1}{2}, 1)$ such that $|P \cap [\frac{1}{2}, \frac{1}{2} + \frac{1}{2^\ell})| = |N \cap [\frac{1}{2}, \frac{1}{2} + \frac{1}{2^\ell})| = 1/2^{\ell+1}$ for each integer $\ell \geq 2$. Take $I_k(x) = (x - \frac{1}{2^k}, x]$ for each $x \in N$, and $I_k(x) = (x, x + \frac{1}{2^k}]$ for each $x \in P$. Let $f(x) = \chi_{[\frac{1}{2},1)}(x)$, and let $\ell > 2$ denote a positive integer. For each $k \geq 2$, and all $x \in [\frac{1}{2}, \frac{1}{2} + \frac{1}{2^\ell})$ each $f_k(x) = 1$. If we also have $x \in P$, then $D_k f(x) = 1$, and we have $D_k f(x) - f_k(x) = 0$. For $k = 1$, we have $f_1(x) = 1$ and $D_1 f(x) = 2(1-x)$, and for $k = 0$ we have $f_0(x) = \frac{1}{2} = D_0 f(x)$. Consequently,

$$Sf(x) = \left(\sum_{k=0}^{\infty} |D_k f(x) - f_k(x)|^2 \right)^{1/2} = 2x - 1 \leq 1.$$

For each k, $2 \leq k < \ell$, if $x \in N \cap [\frac{1}{2}, \frac{1}{2} + \frac{1}{2^\ell})$, then $D_k f(x) = 2^k(x - \frac{1}{2}) \leq 2^k/2^\ell \leq \frac{1}{2}$, and we have $|D_k f(x) - f_k(x)| \geq |\frac{1}{2} - 1| = \frac{1}{2}$. Consequently, for such x we have $Sf(x) \geq \frac{1}{2}\sqrt{\ell - 2}$.

From this we see that if we take $I_\ell = (\frac{1}{2}, \frac{1}{2} + \frac{1}{2^\ell})$, and we compute the average

$$\frac{1}{|I_\ell|^2} \int_{I_\ell} \int_{I_\ell} |Sf(y) - Sf(x)|\, dy\, dx,$$

we see that for ℓ large enough, this average is at least

$$2^{2\ell} \int_{N\cap I_\ell} \int_{P \cap I_\ell} |\tfrac{1}{2}\sqrt{\ell - 2} - 1|\, dy\, dx \geq \tfrac{1}{16}\sqrt{\ell}.$$

Since ℓ can be taken as large as we want, we see that the square function cannot map L^∞ into BMO.

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

Inequalities for L^p, $1 < p < \infty$.

THEOREM 4.6. *Let A be any of the operators A_1, A_2, A_3, A_4 defined in (4.5) and (4.6). Then A is a type (p, p) operator for any $1 < p < \infty$. In case $A = A_1$ and $A = A_2$, the $\ell^p \to \ell^p$ operator bound does not depend on the increasing sequence n_k.*

The same is true in the general ergodic theoretical setting.

Remark. When $A = A_1$, the above theorem improves a result in [**20**] where it was shown that A_1 is type (p, p) for $1 < p \leq 2$.

Proof. In case $A = A_1, A_2, A_3$, we can prove the theorem by interpolation between weak $(1, 1)$ and $\ell^\infty \to BMO$. The necessary interpolation theorem is proved in [**3**].

As for the case $A = A_4$, we cannot directly use the result as it is stated in [**3**]. However, an examination of the proof given there shows their result remains true using only dyadic BMO. In particular, the decomposition in their Lemma 3.2 can be made with dyadic intervals. Then the result follows exactly as in their proof. □

In light of Remark 4, the following question remains.

Question 4.7. Let f_k be the usual dyadic martingale on $[0, 1)$ and let $D_k f(x) = 2^k \int_{I_k(x)} f(t)\,dt$, where $I_k(x)$ is a measurably chosen interval of length $1/2^k$ which contains the point x. Does the square function

$$Sf(x) = \left(\sum_{k=0}^{\infty} |D_k f(x) - f_k(x)|^2 \right)^{1/2}$$

map L^p to L^p for all $p, 1 < p < \infty$?

We have already shown that S maps L^2 to L^2 (see Theorem 2.2). From this, and Theorem 3.1, it is not hard to show that it maps L^1 to weak L^1, so we have the operator bounded from L^p to L^p for $1 < p \leq 2$. Thus the question is really what happens for $2 < p < \infty$?

5. *Square functions and upcrossing results*

We first introduce some notation and give the ergodic theoretical analog of Theorem 6.8. For any sequence of operators, (T_t) and $\lambda > 0$ we define $\Lambda(\lambda, T_t, f, x)$ to be the number of λ jumps of the sequence $(T_t f(x))$, that is, the maximal N for which there are $t_1 < t_2 < \cdots < t_N$ so that

$$|T_{t_m} f(x) - T_{t_{m+1}} f(x)| > \lambda.$$

Let R_t be a nested sequence of intervals in $\mathbb{Z}$ all containing 0 and let $M_t f(x) = (1/|R_t|) \sum_{j \in R_t} f(x + j)$. In what follows, we will often abbreviate $\Lambda(\lambda, M_t, f, x)$ by $\Lambda(f, x)$.

THEOREM 5.1. *There is a constant C so that for any $f \in \ell^1$, $\lambda > 0$ and positive integer n, we have*

$$|\{x : \Lambda(\lambda, f, x) > n\}| \leq \frac{C}{\lambda \sqrt{n}} \|f\|_{\ell^1}.$$

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

Proof. All we need to prove is

$$|\{\lambda\sqrt{\Lambda(\lambda, f, x)} > \gamma\}| \leq \frac{C}{\gamma}\|f\|_{\ell^1}. \tag{5.1}$$

Indeed, take $\gamma = \lambda\sqrt{n}$ to obtain the inequality of our theorem.

In proving (5.1), we can assume that $|R_t| = t$ and that $\gamma = 1$. Let E_t be as in Theorem 2.17. Let us estimate

$$\lambda\sqrt{\Lambda(\lambda, M_t, f, x)} \leq \lambda\sqrt{\Lambda(\lambda/2, M_t - E_t, f, x)} + \lambda\sqrt{\Lambda(\lambda/2, E_t, f, x)}.$$

We estimate the first term on the right-hand side by $\|M_t f(x) - E_t f(x)\|_{v(2)}$. Hence, by Theorem 3.7, we have the inequality

$$\begin{aligned} |\{\lambda\sqrt{\Lambda(\lambda/2, M_t - E_t, f, x)} > 1\}| &= |\{\lambda/2 : \sqrt{\Lambda(\lambda/2, M_t - E_t, f, x)} > 1/2\}| \\ &\leq C\|f\|_{\ell^1}. \end{aligned}$$

That the second term satisfies a weak $(1, 1)$ estimate follows from the inequality in (6.3). □

Of course, this theorem has the following corollary for upcrossings.

First recall the definition of upcrossings. Fix $\beta > \alpha$. Let $N(f, x) = N(\alpha, \beta, f, x)$ be the number of upcrossings of the sequence $(M_t f(x))$, that is, the maximal N for which there are $s_1 < t_1 < s_2 < t_2 < \cdots < s_N < t_N$ so that $M_{s_i} f(x) < \alpha$ and $M_{t_i} f(x) > \beta$ for $i = 1, 2, \ldots, N$.

THEOREM 5.2. *There is a constant C so that for any $f \in \ell^1$, $\beta > \alpha$ and positive integer n, we have*

$$|\{x : N(\alpha, \beta, f, x) > n\}| \leq \frac{C}{(\beta - \alpha)\sqrt{n}}\|f\|_1. \tag{5.2}$$

Remark. In the case $R_n = [0, n)$, Bishop [**5**] (see also [**21**]) proves a stronger result. He proves

$$m\{x : N(\alpha, \beta, f, x) > n\} \leq \frac{1}{(\beta - \alpha)n}\|f\|_1.$$

In the case $R_n = (-n, n)$, Kalikow and Weiss [**21**] prove

$$m\{x : N(\alpha, \beta, f, x) > n\} \leq \frac{C}{\beta - \alpha}\sqrt{\frac{\log n}{n}}\|f\|_1.$$

Our result improves on this inequality in three ways: we improve the constant $\sqrt{\log n/n}$ to $1/\sqrt{n}$; the intervals R_n are more general; finally, we in fact have an inequality for λ-jumps not just for upcrossings. But the fact that we derive our upcrossings result from a λ-jump result possibly prevents us from obtaining the constant $1/n$ in Bishop's result.

Next, we turn to the ergodic theoretical analog of the exponential estimate of Theorem 6.9.

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

Remark. The fact that $\mathcal{V}_\varrho(M_t f(x))$ maps L^∞ to BMO, and consequently we have the John–Nirenberg inequality, gives an exponential estimate on $\mathcal{V}_\varrho(M_t f(x))$, but we are not sure how this relates to the exponential estimate given by Kalikow and Weiss [**21**]. However, by using the information obtained from the square function, as stated in Theorem 5.2 and a covering argument (quite different from that given by Kalikow and Weiss), we can prove an exponential estimate on the number of upcrossings. We note that one advantage of our argument is that we only need to assume the intervals R_t are nested. In particular, we do not assume, as they do, that the R_t satisfy for some $\gamma > 0$, $|R_t \cap \mathbb{Z}^+|/|R_t| > \gamma$. On the other hand, we do not obtain exponential estimates for the upcrossings for a non-negative f—only for bounded f.

It is also interesting to point out that our method of proof of the exponential inequality for ergodic theoretical upcrossings and λ-jumps is completely different from that of the proof of Theorem 6.9. The reason for this is that we could not prove as good a bound for the ergodic theoretical square function type operators as we have for the L^p norm of the martingale square function type operators. Indeed, while in the martingale case the L^p bound goes to infinity as $\sqrt{p}$ (cf. Theorem 6.4), we can presently only prove a growth rate of p for the ergodic theoretical case.

The point of the following is that a non-trivial estimate for the upcrossings for bounded functions is 'self-improving', and implies an exponential estimate for upcrossings for bounded functions. Although we apply the result only for functions defined on $\mathbb{Z}$, we state it for $\mathbb{Z}^d$, because this higher-dimensional result is not harder to prove.

Let $R_1, R_2, \ldots$ be a sequence of rectangles in $\mathbb{Z}^d$. (For our discussion, a rectangle is *always* assumed to have sides parallel to the axes.) For a function $\phi : \mathbb{Z}^d \to \mathbb{R}$ denote

$$M_t\phi(x) = \frac{1}{|R_t|}\sum_{j\in R_t}\phi(x+j).$$

Depending on the context, we will use $|\cdot|$ as usual, to denote the absolute value of a number, but also to denote the cardinality of a set.

For a function $\phi : \mathbb{Z}^d \to \mathbb{R}$ and numbers α, β with $\alpha < \beta$, let $N(\phi, x) = N(\alpha, \beta, \phi, x)$ be the number of upcrossings of the sequence $(M_t\phi(x))$, that is, the maximal N for which there are $s_1, s_2, \ldots, s_N$ and $t_1, t_2, \ldots, t_N$ with

$$s_1 < t_1 < s_2 < t_2 < \cdots < s_N < t_N$$

so that

$$M_{s_1}\phi(x) < \alpha, \quad M_{t_1}\phi(x) > \beta, \quad M_{s_2}\phi(x) < \alpha, \quad \ldots \quad M_{t_N}\phi(x) > \beta.$$

THEOREM 5.3. *Let $R_1, R_2, \ldots$ be a sequence of rectangles in $\mathbb{Z}^d$ so that each rectangle contains the origin and the sequence of sidelengths in each coordinate direction is non-decreasing (for example, a nested sequence of rectangles). Suppose that for every pair (α, β) of numbers with $\alpha < \beta$, there is a function $\epsilon(q) = \epsilon(\alpha, \beta, q)$ so that:*

(i) $\lim_{q\to\infty}\epsilon(q) = 0$;

(ii) *for every function $\phi : \mathbb{Z}^d \to \mathbb{R}$ of finite support and $\|\phi\|_{\ell^\infty(Z^d)} \le 1$, we have*

$$|\{x \mid N(\phi, x) \ge q\}| \le \epsilon(q)\cdot\|\phi\|_{\ell^1(Z^d)}.$$

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

Then for every pair (α, β) of numbers with $\alpha < \beta$, there are positive constants $C = C(\alpha, \beta)$ and $c = c(\alpha, \beta)$, with $c < 1$, so that for any dynamical system (X, Σ, m), and function $f : X \to \mathbb{R}$, with $\|f\|_{L^\infty(X)} \le 1$, we have

$$m\{x \mid N(f, x) \ge n\} \le C \cdot c^n \cdot \|f\|_{L^1(X)}.$$

Remark. It is easy to construct a sequence of intervals so that their lengths form an increasing sequence but, for every q, there is ϕ with $\|\phi\|_{\ell^\infty(\mathbb{Z})} \le 1$ and of finite support so that

$$|\{x \mid N(-1/4, 1/4, \phi, x) \ge q\}| > \tfrac{1}{10} \cdot \|\phi\|_{\ell^1(\mathbb{Z})}. \tag{5.3}$$

However, if the intervals are nested, we have control of the operator $\mathcal{V}_\varrho(M_t f(x))$ and we have seen above that an inequality like (5.3) is not possible.

In fact we have the following.

THEOREM 5.4. *If (R_t) is a nested sequence of intervals in $\mathbb{Z}$, then for all $f \in L^\infty$ with $\|f\|_\infty \le 1$, we have the estimate*

$$m\{x \mid N(f, x) \ge n\} \le C \cdot c^n \cdot \|f\|_{L^1(X)}.$$

Proof. Apply Theorem 5.2 to obtain the non-trivial estimate required in the hypothesis of Theorem 5.3. Then apply Theorem 5.3 and we are done. □

Here is a corollary of Theorem 5.4 the proof of which is almost immediate from the pigeonhole principle. Let $\lambda > 0$, and let $\Lambda(f, x) = \Lambda(\lambda, f, x)$ denote the number of λ jumps of the sequence $(M_t f(x))$, that is, the maximal N for which there are $t_1 < t_2 < \cdots < t_N$ so that

$$|M_{t_m} f(x) - M_{t_{m+1}} f(x)| > \lambda.$$

COROLLARY 5.5. *Let $\lambda > 0$. With the assumptions of Theorem 5.3, there are positive constants $C = C(\lambda)$ and $c = c(\lambda)$, with $c < 1$, so that for any dynamical system (X, Σ, m), and function $f : X \to \mathbb{R}$, with $\|f\|_{L^\infty(X)} \le 1$, we have*

$$m\{x \mid \Lambda(f, x) \ge n\} \le C \cdot c^n \cdot \|f\|_{L^1(X)}.$$

Proof of Theorem 5.3. By the transference principle, it is enough to prove the inequality on $\mathbb{Z}^d$. So we will prove the inequality

$$|\{x \mid N(\phi, x) \ge n\}| \le C \cdot c^n \cdot \|\phi\|_{\ell^1(\mathbb{Z}^d)}.$$

In fact, we give the proof in case $d = 2$ only; there is no further idea involved in proving it for arbitrary d. We may also assume that $\alpha > 0$.

We need to start with a lemma, which requires some notation. For each q let us define the following two sets:

$$\mathcal{F}_q = \mathcal{F}_q(\alpha, \beta) = \mathcal{F}_q(\alpha, \beta, \phi) = \{x \mid N(\phi, x) \ge q\};$$

and let $\mathcal{G}_q = \mathcal{G}_q(\alpha, \beta, \phi)$ contain those rectangles G which have the property that for some x and $s_1 < t_1 < s_2 < t_2 < \cdots < s_q < t_q$ we have that $G = x + R_{s_1}$ and

$$M_{s_m}\phi(x) < \alpha, \quad M_{t_m}\phi(x) > \beta.$$

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

Since each R_s contains the origin, it is clear that $\mathcal{F}_q \subseteq \cup\mathcal{G}_q$. The assumptions (i) and (ii) in our theorem say that $\mathcal{F}_q$ gets smaller as q gets large. The content of our lemma is that the same happens with $\cup\mathcal{G}_q$.

LEMMA 5.6. *For every pair (α, β) of numbers with $\alpha < \beta$, there is a function $\eta(q) = \eta(\alpha, \beta, q)$ so that:*
(iii) $\lim_{q\to\infty} \eta(q) = 0$*;*
(iv) *for every function $\phi : \mathbb{Z}^d \to \mathbb{R}$ of finite support and $\|\phi\|_{\ell^\infty(\mathbb{Z}^d)} \leq 1$, we have*

$$|\cup \mathcal{G}_q| \leq \eta(q) \cdot \|\phi\|_{\ell^1(\mathbb{Z}^d)}.$$

Proof of Lemma 5.6. By the usual procedure (at each step selecting rectangles of maximal size), we can select $G_1, G_2, \ldots$, a collection of pairwise disjoint rectangles, from $\mathcal{G}_q$ so that

$$|\cup_i G_i| = \sum_i |G_i| \geq \frac{1}{3^2} \cdot |\cup \mathcal{G}_q|. \tag{5.4}$$

Let $\alpha' = \alpha + (\beta - \alpha)/3$ and $\beta' = \beta - (\beta - \alpha)/3$. All we have to show is that

$$|\cup_i G_i| \leq \left(\frac{12}{\beta - \alpha}\right)^2 \cdot |\mathcal{F}_q(\alpha', \beta')|. \tag{5.5}$$

Indeed, this implies, together with (5.4), that

$$|\cup \mathcal{G}_q| \leq \left(\frac{36}{\beta - \alpha}\right)^2 \cdot |\mathcal{F}_q(\alpha', \beta')|.$$

Since, by assumption (ii), we have

$$|\mathcal{F}_q(\alpha', \beta')| \leq \epsilon(q, \alpha', \beta') \cdot \|f\|_{\ell^1},$$

we see we can put

$$\eta(q, \alpha, \beta) = \epsilon(q, \alpha', \beta') \cdot \left(\frac{36}{\beta - \alpha}\right)^2.$$

To prove (5.5) let $x_i \in \mathcal{F}_q(\alpha, \beta) \cap G_i$, and let H_i be a rectangle which contains x_i, H_i is contained in G_i, and the length of the sides of H_i are $(\beta - \alpha)/12$ times the length of the corresponding sides of G_i. Our purpose is to prove that, for each i, we have

$$H_i \subseteq \mathcal{F}_q(\alpha', \beta'). \tag{5.6}$$

Note that

$$|H_i| = \left(\frac{\beta - \alpha}{12}\right)^2 \cdot |G_i|.$$

Summing this over i we get, using (5.6) and the fact that both the G_i's and the H_i's are disjoint,

$$\left(\frac{\beta - \alpha}{12}\right)^2 \cdot |\cup_i G_i| = |\cup_i H_i| \leq |\mathcal{F}_q(\alpha', \beta')|,$$

proving (5.5).

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

In proving (5.6), we simplify our notation by suppressing the index i. Since $x \in \mathcal{F}_q$, there are $s_1 < t_1 < s_2 < t_2 < \cdots < s_q < t_q$ such that

$$M_{s_m}\phi(x) < \alpha, \quad M_{t_m}\phi(x) > \beta.$$

Let $y \in H$. Our proof is finished if we show $M_{s_m}\phi(y) < \alpha'$, and $M_{t_m}\phi(y) > \beta'$ for every $m \le q$.

If R is a rectangle which contains x and its sidelengths are at least as large as the corresponding sidelengths of G, then we have, as simple geometry shows,

$$|(x+R)\Delta(y+R)| = |R\Delta(y-x+R)| \le \frac{\beta-\alpha}{3}\cdot|R|.$$

(In the above, Δ denotes symmetric difference.) This and $\|\phi\|_{\ell^\infty} \le 1$ imply that

$$\left|\sum_{j\in y+R}\phi(j) - \sum_{j\in x+R}\phi(j)\right| = \left|\sum_{j\in(x+R)\Delta(y+R)}\phi(j)\right| \le \frac{\beta-\alpha}{3}\cdot|R|.$$

But this estimate implies both $M_{s_m}\phi(y) < \alpha'$ and $M_{t_m}\phi(y) > \beta'$ for $m \le q$. Indeed, take $R = R_{s_m}$ and $R = R_{t_m}$, respectively, and recall that

$$\sum_{j\in x+R_{s_m}}\phi(j) < \alpha|R_{s_m}|, \qquad \sum_{j\in x+R_{t_m}}\phi(j) > \beta|R_{t_m}|.$$

□

Let us turn to the proof of the theorem. We shall prove that, for any positive integers p and q, we have

$$|\cup\mathcal{G}_{(p+1)q}| \le \eta(q)\cdot 5^2\cdot|\cup\mathcal{G}_{pq}|. \tag{5.7}$$

How does this prove the theorem? First, choose q so large that $\eta(q)\cdot 5^2$ is less than 1, say, $\eta(q)\cdot 5^2 \le 1/2$. We can choose such a q by the lemma. Let us assume that $n \ge q$, and let us write $n = sq + r$, with $0 \le r < q$. By (5.7), we have

$$|\mathcal{F}_n| \le |\mathcal{F}_{sq}| \le |\cup\mathcal{G}_{sq}| \le (\tfrac{1}{2})^{s-1}\cdot|\cup\mathcal{G}_q|.$$

But, by our lemma again, we have

$$|\cup\mathcal{G}_q| \le \eta(q)\cdot\|\phi\|_{\ell^1} \le \tfrac{1}{2}\cdot\|\phi\|_{\ell^1}.$$

It follows, since $s = (n-r)/q$, that

$$|\mathcal{F}_n| \le (\tfrac{1}{2})^s\cdot\|\phi\|_{\ell^1} = (\tfrac{1}{2})^{-r/q}\cdot(\sqrt[q]{1/2})^n\cdot\|\phi\|_{\ell^1},$$

and hence, if we choose $C = 2$ and $c = \sqrt[q]{1/2}$, we have the theorem for $n \ge q$. We can take care of the rest of the n, that is $n < q$, if we set

$$C = 2 + \max_{n<q}\frac{\eta(n)}{c^n}.$$

So let us prove (5.7). First, to every rectangle $H \in \mathcal{G}_{(p+1)q}$ we correspond a rectangle $H^+ \in \mathcal{G}_{pq}$ in the following way. Since $H \in \mathcal{G}_{(p+1)q}$, for some x and $s_1 < t_1 < s_2 < t_2 < \cdots < s_{(p+1)q} < t_{(p+1)q}$ we have that $H = x + R_{s_1}$ and

$$M_{s_m}\phi(x) < \alpha, \quad M_{t_m}\phi(x) > \beta.$$

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

By the definition of $\mathcal{G}_{pq}$, we have $x + R_{s_{q+1}} \in \mathcal{G}_{pq}$, and we take $H^+ = x + R_{s_{q+1}}$.

Let us denote $\mathcal{H} = \mathcal{G}_{(p+1)q}$ and

$$\mathcal{K} = \left\{K \mid K \in \mathcal{G}_{pq};\, K = H^+ \text{ for some } H \in \mathcal{H}\right\}.$$

Since $\mathcal{K} \subseteq \mathcal{G}_{pq}$, the inequality in (5.7) follows from

$$|\cup \mathcal{H}| \leq \eta(q) \cdot 5^2 \cdot |\cup \mathcal{K}|.$$

Technically, this inequality is proved in the following form. We are going to select $K_1, K_2, \ldots$, a collection of pairwise disjoint rectangles from $\mathcal{K}$, and we will prove that

$$|\cup \mathcal{H}| \leq \eta(q) \cdot 5^2 \cdot |\cup_i K_i|. \tag{5.8}$$

Although the proof of (5.8) requires some new notation, it is not much more than the usual greedy selection.

Let $K_1 \in \mathcal{K}$ be a rectangle of maximal size. Let us set $\mathcal{H}_1 = \mathcal{H}$, $\mathcal{K}_1 = \mathcal{K}$ and define the sets

$$\begin{aligned} \mathcal{I}_1 &= \{H \mid H \in \mathcal{H}_1, H^+ \cap K_1 \neq \emptyset\}, \\ \mathcal{L}_1 &= \{K \mid K \in \mathcal{K}_1, K \cap K_1 \neq \emptyset\}. \end{aligned}$$

We assert that

$$|\cup \mathcal{I}_1| \leq \eta(q) \cdot |5K_1|. \tag{5.9}$$

Indeed, for $H \in \mathcal{I}_1$ there are x and $s_1 < t_1 < s_2 < t_2 < \cdots < s_{(p+1)q} < t_{(p+1)q}$ such that $H = x + R_{s_1}$, $H^+ = x + R_{s_{q+1}}$ and

$$M_{s_m}\phi(x) < \alpha, \quad M_{t_m}\phi(x) > \beta.$$

We observe that, as a consequence of the maximal size of K_1, for $n \leq q+1$, we have $x + R_{s_n} \subseteq 5 \cdot K_1$, where, for a rectangle K, $5K$ denotes the rectangle with the same center as K and its sidelengths are 5 times that of K. It follows that $H \in \mathcal{G}_q(\phi \cdot \chi_{5K_1})$. As a consequence, we have

$$\mathcal{I}_1 \subseteq \mathcal{G}_q(\phi \cdot \chi_{5K_1}).$$

But now

$$|\cup G_q(\phi \cdot \chi_{5K_1})| \leq \eta(q) \cdot \|f \cdot \chi_{5K_1}\|_{\ell^1} \leq \eta(q) \cdot \|\chi_{5K_1}\|_{\ell^1},$$

where $\|f\|_{\ell^\infty} \leq 1$ is used; and (5.9) follows.

A consequence of (5.9) is

$$|\cup \mathcal{I}_1| \leq \eta(q) \cdot 5^2 \cdot |K_1|.$$

Let $\mathcal{H}_2 = \mathcal{H}_1 \setminus \mathcal{I}_1$ and $\mathcal{K}_2 = \mathcal{K}_1 \setminus \mathcal{L}_1$. Select $K_2 \in \mathcal{K}_2$ of maximal size. Then K_2 is disjoint from K_1. Define the sets

$$\begin{aligned} \mathcal{I}_2 &= \{H \mid H \in \mathcal{H}_2, H^+ \cap K_2 \neq \emptyset\}, \\ \mathcal{L}_2 &= \{K \mid K \in \mathcal{K}_2, K \cap K_2 \neq \emptyset\}. \end{aligned}$$

Similar to the first step of the iteration, we obtain

$$|\cup \mathcal{I}_2| \leq \eta(q) \cdot |5K_2|,$$

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

and hence

$$|\cup \mathcal{I}_2| \le \eta(q) \cdot 5^2 \cdot |K_2|.$$

We can repeat this process only finitely many times. Then we obtain the pairwise disjoint rectangles $K_1, K_2, \ldots$, and the sets $\mathcal{I}_1, \mathcal{I}_2, \ldots$, $\mathcal{L}_1, \mathcal{L}_2, \ldots$ with the properties

$$\cup_i \mathcal{L}_i = \mathcal{K}, \quad \cup_i \mathcal{I}_i = \mathcal{H}, \quad |\cup \mathcal{I}_i| \le \eta(q) \cdot 5^2 \cdot |K_i|.$$

Summing this latter inequality over i and using that the K_i's are pairwise disjoint, we get

$$|\cup \mathcal{H}| \le \sum_i |\cup \mathcal{I}_i| \le \sum_i \eta(q) \cdot 5^2 \cdot |K_i| = \eta(q) \cdot 5^2 \cdot |\cup_i K_i|,$$

which is (5.8). □

Remark.

- It is only in Lemma 5.6 that something special about the geometry of $\mathbb{Z}^2$ is used. The rest of the proof works with 'Tempelman-type' assumptions in any locally compact topological group.
- The boundedness of ϕ was crucially used only in Lemma 5.6.
- Note also, that the only information we need about $\eta(q)$ is that it is less than $1/5^2$ for some q.

6. *Martingale results*

In this section we give the martingale analogs for some of the ergodic theory results discussed earlier in the paper. These martingale results can be read independently of the rest of the paper, and in some cases suggest what the correct ergodic theory result might be.

Inequalities for L^p, $1 \le p < \infty$. Let (X, Σ, m) denote a probability space and $(\mathcal{F}_n)$ a non-decreasing sequence of sub-σ-fields of Σ. We will denote by $f = (f_1, f_2, \ldots)$ a martingale or submartingale relative to the sequence $(\mathcal{F}_n)$. For $0 < p < \infty$ we let $\|f\|_p = \sup_k \|f_k\|_p$. As usual, $f_n^\star$ denotes the maximal function, $f_n^\star = \max_{1 \le k \le n} |f_k|$. We begin with the following simple theorem, the reverse martingale part of which was used in the proof of Theorem 2.6. For a sequence $n_1 < n_2 < \cdots$ of positive integers define

$$\begin{aligned} Of &= \mathcal{O}(f_n) \\ &= \left(\sum_{k=1}^{\infty} \sup_{n_{k-1} \le n \le m < n_k} |f_n - f_m|^2 \right)^{1/2}. \end{aligned} \tag{6.1}$$

THEOREM 6.1. *Let $f = (f_n)$ be a martingale. Then there is a constant c so that for any sequence $n_1 < n_2 \cdots$ of positive integers we have*

$$\begin{aligned} \|Of\|_2 &\le c\|\mathcal{S}(f_{n_k} - f_{n_{k-1}})\|_2 \\ &\le c\|f\|_2, \end{aligned}$$

where Of is defined in (6.1).

The same conclusion holds if f_n is a reverse martingale.

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

Proof. We just give the proof when f_n is a martingale; the proof of the reverse martingale case is the same.

By (1.1), it is enough to estimate

$$\begin{aligned} O'f &= \mathcal{O}'(f_n) \\ &= \left(\sum_{k=1}^{\infty} \sup_{n_{k-1} \le n < n_k} |f_n - f_{n_{k-1}}|^2 \right)^{1/2}. \end{aligned}$$

First note that

$$\left\| \sup_{n_{k-1} < n < n_k} |f_n - f_{n_{k-1}}| \right\|_2 \le c \| f_{n_k} - f_{n_{k-1}} \|_2.$$

This is a consequence of Doob's maximal theorem applied to the martingale $X_n = f_{n_{k-1}+n} - f_{n_{k-1}}$, where $0 \le n \le n_k - n_{k-1}$. That is, $\|X^\star\|_2 \le c\|X_{n_{k-1}-n_k}\|_2 \le c\|f_{n_k} - f_{n_{k-1}}\|_2$. Hence if (d_k) denotes the difference sequence, i.e. $d_k = f_k - f_{k-1}$, then using the orthogonality of (d_j), we have

$$\begin{aligned} \|O'f\|_2^2 &\le \sum_{k=1}^{\infty} c^2 \|f_{n_k} - f_{n_{k-1}}\|_2^2 \\ &\le \sum_{k=1}^{\infty} \left\| \sum_{j=n_{k-1}+1}^{n_k} d_j \right\|_2^2 \\ &\le \sum_{j=1}^{\infty} c^2 \|d_j\|_2^2 \\ &\le c^2 \|f\|_2^2. \end{aligned}$$

□

The following two theorems and their proofs were pointed out to us by D. Burkholder. The proofs are similar to those given by Burkholder [**9**] to prove that the standard martingale square function satisfies similar inequalities.

Let us define the operator $V_\varrho f$ by

$$\begin{aligned} V_\varrho f &= \mathcal{V}_\varrho(f_n) \\ &= \sup_{(n(i))} \left(\sum_{k=1}^{\infty} |f_{n(i)} - f_{n(i+1)}|^\varrho \right)^{1/\varrho}, \end{aligned} \tag{6.2}$$

where the supremum is taken over all sequences $n(1) < n(2) < \cdots$.

THEOREM 6.2. *Let $f = (f_n)$ be a submartingale.*

(a) *For $\varrho > 2$, the operator $V_\varrho f$ defined in (6.2) is weak type* $(1, 1)$.

(b) *For any fixed increasing sequence of non-negative integers, the operator Of defined in (6.1) is weak type* $(1, 1)$.

Proof. To simplify notation, we set $V = V_\varrho$.

We first handle the martingale case. In this case the weak type $(1, 1)$ result for both operators will follow from the martingale analog of the Calderón–Zygmund decomposition. This decomposition, known as Gundy's decomposition [**18**], is stated as the following lemma. (See also [**9**].)

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

LEMMA 6.3. *Suppose $f = (f_1, f_2, \ldots)$ is an L^1 bounded martingale and λ is a positive real number. Then there are martingales X, Y, Z, with corresponding difference sequences x, y, z all relative to $\mathcal{F}_n$ such that:*
(i) $f = X + Y + Z$;
(ii) $\|X\|_2^2 \leq 2\lambda\|f\|_1$;
(iii) $\|\sum_{k=1}^{\infty} |y_k|\|_1 \leq 4\|f\|_1$;
(iv) $P(Z^\star > 0) \leq \|f\|_1/\lambda$.

Using this lemma, we first show that Vf is a weak type $(1, 1)$ operator. Since V is sublinear, we have

$$P(Vf > \lambda) \leq P\left(VX > \frac{\lambda}{3}\right) + P\left(VY > \frac{\lambda}{3}\right) + P\left(VZ > \frac{\lambda}{3}\right).$$

For the first term we use the known L^2 boundedness of the operator V (Lepingle's theorem) and property (ii) to obtain

$$\begin{aligned} P\left(VX > \frac{\lambda}{3}\right) &\leq \frac{9}{\lambda^2}\|VX\|_2^2 \\ &\leq \frac{c(\varrho)}{\lambda^2}\|X\|_2^2 \\ &\leq \frac{c(\varrho)}{\lambda^2} 2\lambda\|f\|_1 \\ &\leq \frac{c(\varrho)}{\lambda}\|f\|_1. \end{aligned}$$

For the second piece, since $VY \leq \sum_k |y_k|$, we have

$$\begin{aligned} P\left(VY > \frac{\lambda}{3}\right) &\leq \frac{3}{\lambda}\|VY\|_1 \\ &\leq \frac{3}{\lambda}\left\|\sum_{k=1}^{\infty} |y_k|\right\|_1 \\ &\leq \frac{c}{\lambda}\|f\|_1. \end{aligned}$$

For the final piece we have $P(VZ > \lambda/3) \leq P(Z^\star > 0) \leq (1/\lambda)\|f\|_1$ and we are done with V. The proof for O follows in the same way. The necessary L^2 bound on O is from Theorem 6.1 above.

In the submartingale case we note that if (f_n) is an L_1 bounded submartingale, then we can write $f_n = m_n + I_n$, where m_n is an L^1 bounded martingale and I_n is an increasing process with $I_0 = 0$. We have

$$I_n = I_{n-1} + E(f_n - f_{n-1}|\mathcal{F}_{n-1}).$$

Hence

$$\begin{aligned} E(I_n) &= E(E(f_n - f_{n-1}|\mathcal{F}_{n-1})) + E(I_{n-1}) \\ &= E(f_n) - E(f_{n-1}) + E(I_{n-1}). \end{aligned}$$

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

By induction, we see that $E(I_n) = E(f_n) - E(f_0)$. Hence $\|I_n\|_1 \leq 2\|f\|_1$. This also implies $\|m_n\|_1 \leq 3\|f\|_1$.

We have

$$m\{Vf > \lambda\} \leq m\left\{Vm > \frac{\lambda}{2}\right\} + m\left\{VI > \frac{\lambda}{2}\right\}.$$

For the first piece we use the martingale result above. For the second, note that

$$\begin{aligned} VI &= \sup_{(n_i)} \left(\sum_{i=1}^{\infty} |I_{n_i} - I_{n_{i-1}}|^{\varrho} \right)^{1/\varrho} \\ &\leq \sup_{(n_i)} \sum_{i=1}^{\infty} |I_{n_i} - I_{n_{i-1}}| \\ &\leq \sum_{i=1}^{\infty} I_i - I_{i-1} = \lim_{i\to\infty} I_i. \end{aligned}$$

From this we see that

$$m\left\{VI > \frac{\lambda}{2}\right\} \leq m\left\{\lim_n I_n > \frac{\lambda}{2}\right\} \leq \frac{2}{\lambda}\left\|\lim_n I_n\right\|_1 \leq \frac{c}{\lambda}\sup_n \|f_n\|_1. \qquad \square$$

We also have the following.

THEOREM 6.4. *Let $f = (f_1, f_2, \ldots)$ be a martingale or a non-negative submartingale, $1 < p < \infty$ and $1/p + 1/q = 1$.*

(i) *For $\varrho > 2$ there is a positive real number $c(\varrho)$ such that*

$$\|V_{\varrho} f\|_p \leq c(\varrho) p^{1/\varrho} q \|f\|_p.$$

(ii) *There is a positive real number c such that*

$$\|Of\|_p \leq c p^{1/2} q \|f\|_p.$$

Proof. We will prove the Theorem only for $V = V_{\varrho}$, the proof for O being very similar.

We define the operator

$$V^n f = V^n_{\varrho} f = \sup_{(n_i)\leq n} \left(\sum_i |f_{n_i} - f_{n_{i-1}}|^{\varrho} \right)^{1/\varrho},$$

where $\sup_{(n_i)\leq n}$ denotes taking the supremum over all increasing sequences of non-negative integers, with largest entry in the sequence less than or equal to n. Clearly, $V^n f$ is a non-decreasing sequence since for later entries in the sequence we are taking a supremum over a larger set. Also define the operator

$$O^n f = \left(\sum_{n_k \leq n} \sup_{n_{k-1} \leq \ell < m < n_k} |f_{\ell} - f_m|^2 \right)^{1/2}$$

in a similar way.

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

LEMMA 6.5. *Let $f = (f_1, f_2, \ldots)$ be a non-negative submartingale and n a positive integer.*

(i) *If $Y = V^n(\theta f) \vee f_n^\star$ then*

$$\lambda P(Y > \beta\lambda) \leq c(\varrho) \int_{\{Y>\lambda\}} f_n,$$

where $\theta > 0$ and $\beta = (1 + 2\theta^\varrho)^{1/\varrho}$.

(ii) *If $Y = O^n(\theta f) \vee f_n^\star$ then*

$$\lambda P(Y > \beta\lambda) \leq c(\varrho) \int_{\{Y>\lambda\}} f_n,$$

where $\theta > 0$ and $\beta = (1 + 2\theta^2)^{1/2}$.

Remark. The role of θ in the above lemma is to help obtain the rate of growth of the L^p norm of V_ϱ that is given in Theorem 6.4. We will later see that the optimal choice of θ is $(2p-2)^{1/\varrho}$. The role of $f_n^\star$ is to control the size of the jumps of the random variable Y.

Proof. Since $\beta > 1$ we see that

$$\lambda P(Y > \beta\lambda) \leq \lambda P(f_n^\star > \lambda) + \lambda P(V^n(\theta f) > \beta\lambda,\, f_n^\star \leq \lambda).$$

The first term on the right-hand side is dominated by $\int_{\{f_n^\star>\lambda\}} f_n \leq \int_{\{Y>\lambda\}} f_n$. Hence it is enough to consider the set

$$B = \{V^n(\theta f) > \beta\lambda,\, f_n^\star \leq \lambda\}.$$

Let $g_k = \chi_{\{V^k\theta f>\lambda\}} f_k = I_k f_k$. Note that $I_k \leq I_{k+1}$ and hence

$$\begin{aligned} E(g_{k+1}|\mathcal{F}_k) &= E(I_{k+1}f_{k+1}|\mathcal{F}_k) \\ &\geq I_k E(f_{k+1}|\mathcal{F}_k) \\ &\geq I_k f_k = g_k. \end{aligned}$$

Hence g_k is a non-negative submartingale.

Now define the stopping time $\tau = \inf\{k \mid V^k(\theta f) > \lambda\}$. For $\omega \in B$ we have $1 \leq \tau \leq n$ and $g_n^\star \leq \lambda$. We also have a sequence $(n_i) \leq n$ such that $\beta^\varrho\lambda^\varrho \leq \sum_i |f_{n_i} - f_{n_{i-1}}|^\varrho$. Let u denote the first value n_i such that $n_i > \tau$ and let ℓ denote the previous value in the sequence (n_i). We then have

$$\begin{aligned} \beta^\varrho\lambda^\varrho &\leq (V^n(\theta f))^\varrho \\ &\leq (V^{\tau-1}(\theta f))^\varrho + \theta^\varrho |f_u - f_\ell|^\varrho + \theta^\varrho \sum_{n_{i-1}>\tau} |f_{n_i} - f_{n_{i-1}}|^\varrho \\ &\leq \lambda^\varrho + \theta^\varrho\lambda^\varrho + \theta^\varrho (V^n g)^\varrho. \end{aligned}$$

Solving, we see that $V^n(g) > \lambda$.

From this we now have

$$\begin{aligned} \lambda P(\{V^n(\theta f) > \beta\lambda,\, f_n^\star \leq \lambda\}) &\leq \lambda P(V^n(g) > \lambda) \\ &\leq c_0(\varrho)\|g\|_1 \\ &\leq c_0(\varrho) \int_{\{Y>\lambda\}} f_n. \end{aligned}$$

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

Hence the Lemma is proved for V_ϱ with $c(\varrho) = c_0(\varrho) + 1$. The same argument also proves the result for O. □

We now continue with the proof of Theorem 6.4. Note that with $1/p + 1/q = 1$ we have the following standard argument:

$$\begin{aligned}
\|Y\|_p^p &= \int_0^\infty p u^{p-1} P(Y > u)\, du \\
&= p \int_0^\infty \beta^p \lambda^{p-1} P(Y > \beta\lambda)\, d\lambda \\
&\le c(\varrho) p \int_0^\infty \beta^p \lambda^{p-2} \int_{\{Y>\lambda\}} f_n \, d\omega \, d\lambda \\
&\le c(\varrho) q \beta^p \int_X f_n \int_0^{Y(\omega)} \lambda^{p-2} \, d\lambda \, d\omega \\
&\le c(\varrho) \frac{1}{q} \beta^p \int_X f_n Y^{p-1} \, d\omega \\
&\le c(\varrho) \frac{1}{q} \beta^p \left(\int_X f_n^p \, d\omega \right)^{1/p} \left(\int_X Y^{(p-1)q} \, d\omega \right)^{1/q} \\
&\le c(\varrho) q \beta^p \|f_n\|_p \|Y\|_p^{p-1}.
\end{aligned}$$

Solving for $\|Y\|_p$ and noting that $\|Vf\|_p \le \|Y\|_p$ completes the proof that $\|Y\|_p \le c(\varrho) q \beta^p \|f_n\|_p$ if f is a non-negative submartingale. Note that since we are using only a finite submartingale in the above argument, we will have $\|Y\|_p < \infty$ and we can solve for $\|Y\|_p$ in the above expression.

Recalling that $Y = V^n(\theta f) \vee f_n^\star$, and recalling the definition of β, we see that

$$\|V_\varrho\|_p \le c(\varrho) q \frac{(1 + 2\theta^\varrho)^{p/\varrho}}{\theta}.$$

Taking the derivative shows the optimal choice of θ is $(2p-2)^{-1/\varrho}$. Using this θ gives the result in the case of a non-negative submartingale.

To complete the proof in the martingale case, we just write $g_n = E(f_n^+ \mid \mathcal{F}_k)$ and $h_n = E(f_n^- \mid \mathcal{F}_k)$. Then $f_n = g_n - h_n$ and $V_n f \le V_n g + V_n h$. Using the above results on the two non-negative submartingales completes the proof.

The same argument handles the case of Of. □

Remark. The L^p boundedness of the operators V_ϱ and O also follows from a result in an earlier paper by Burkholder *et al* [**11**]. It is easy to check that the operators V and O both satisfy the hypothesis of Theorem 2.1 of their paper, and consequently by their theorem, both operators are bounded on all L^p spaces for $1 < p < \infty$. However, because the theorem in [**11**] is quite general, it does not give the order of magnitude of the L^p bound. The order of magnitude given in Theorem 6.4 is probably the best possible. For the standard martingale square function, it is known that the right order is $p^{1/2}q$.

Exponential inequalities. The rate of growth of the L^p norm for the operators V_ϱ and O given by Theorem 6.4 has the following consequence.

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

THEOREM 6.6. *There are positive constants α and K such that for all L^∞ bounded martingales $f = (f_n)$ with $\|f\|_\infty \le 1$ we have*
(a)

$$\int_X e^{\alpha(V_\varrho f)^\varrho}\, d\omega \le K.$$

In this case, K and α depend on ϱ.
(b)

$$\int_X e^{\alpha(Of)^2}\, d\omega \le K.$$

Proof. This is immediate from Theorem 6.4 above and Theorem 4.41 on page 119 of [**28**]. □

Results for λ-jumps. Define $\Lambda_\lambda f(\omega)$ to be the number of λ-jumps of the martingale (f_n). That is, $\Lambda_\lambda f(\omega)$ is the largest integer n such that we can find $0 \le s_1 < t_1 \le s_2 < \cdots \le s_n < t_n$ with the property that $|f_{t_i}(\omega) - f_{s_i}(\omega)| \ge \lambda$ for $i = 1, 2, \ldots, n$.

It is possible to control the λ-jumps of a martingale by the square function of an associated martingale. Recall that if (τ_k) is an increasing sequence of stopping times for a martingale f, and $g_k = f_{\tau_k}$ then (g_k) is a martingale with $\|g\|_p \le \|f\|_p$ for every p, $1 \le p \le \infty$. (See [**14**, p. 300].)

All of our results are easy consequences of the results in the preceding subsections and the following lemma.

LEMMA 6.7. *Let (f_n) be a uniformly integrable martingale sequence. Then there exists an increasing sequence of stopping times (τ_k) such that if $g_k = f_{\tau_k}$ then*

$$\lambda \cdot \sqrt{\Lambda_\lambda f(\omega)} \le 2 \cdot S(g)(\omega).$$

Proof. Fix ω such that $\Lambda_\lambda f(\omega) = n$. Then we can find $0 \le s_1 < t_1 \le s_2 < \cdots \le s_n < t_n$ with the property that $|f_{t_i}(\omega) - f_{s_i}(\omega)| > \lambda$ for $i = 1, 2, \ldots, n$. Set $\tau_0 = 0$ and if τ_{k-1} is defined, define $\tau_k = \inf\{j > \tau_{k-1} \mid |f_j - f_{\tau_{k-1}}| > \frac{1}{2}\lambda\}$. Clearly, either

$$|f_{s_1}(\omega) - f_{\tau_0}(\omega)| > \tfrac{1}{2}\lambda \quad \text{or} \quad |f_{t_1}(\omega) - f_{\tau_0}(\omega)| > \tfrac{1}{2}\lambda.$$

Hence $\tau_1 < t_1$. Next, either

$$|f_{s_2}(\omega) - f_{\tau_1}(\omega)| > \tfrac{1}{2}\lambda \quad \text{or} \quad |f_{t_2}(\omega) - f_{\tau_1}(\omega)| > \tfrac{1}{2}\lambda.$$

We continue inductively, and obtain an increasing sequence of stopping times such that $|f_{\tau_j} - f_{\tau_{j-1}}| > \frac{1}{2}\lambda$ for $j = 1, 2, \ldots, n$.

Letting $g_k = f_{\tau_k}$, we now have $S(g) \ge (\lambda/2)\sqrt{\Lambda_\lambda f}$ which was to be proven. □

THEOREM 6.8. *There is a positive constant C so that*

$$m\{\omega : \Lambda_\lambda f(\omega) > n\} \le \frac{C}{\lambda\sqrt{n}} \cdot \|f\|_1$$

for all integrable martingales $f = (f_n)$ and positive integers n.

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

Proof. The estimate

$$m\{\lambda \cdot \sqrt{\Lambda_\lambda f} > \gamma\} \leq \frac{C}{\gamma}\|f\|_1 \tag{6.3}$$

follows immediately from 6.7 and part (b) of Theorem 6.2 (since $Sg \leq Og$). Now just take $\gamma = \lambda\sqrt{n}$ in (6.3). □

THEOREM 6.9. *There are positive constants α and K such that*

$$m\{\omega : \Lambda_\lambda f(\omega) > n\} \leq Ke^{-\alpha n\lambda^2}$$

for all L^∞ bounded martingales $f = (f_n)$ with $\|f\|_\infty \leq 1$ and for all positive integers n.

Proof. The estimate

$$m\{\lambda^2 \cdot \Lambda_\lambda f > \gamma\} \leq Ke^{-\alpha\gamma} \tag{6.4}$$

follows immediately from Lemma 6.7 and the exponential estimate in part (b) of Theorem 6.6 (since $Sg \leq Og$). Now just take $\gamma = \lambda^2 n$ in (6.4). □

Acknowledgements. We would like to express our thanks to Donald Burkholder for many helpful comments about the paper, and to Yoram Sagher who pointed out reference [**3**] to us.

REFERENCES

[1] M. Akcoglu, A. Bellow, R. L. Jones, V. Losert, K. Reinhold-Larsson and M. Wierdl. The strong sweeping out property for lacunary sequences, Riemann sums, convolution powers, and related matters. *Ergod. Th. & Dynam. Sys.* **16** (1996), 207–253.

[2] I. Assani, K. Petersen and H. White. Some connections between ergodic theory and harmonic analysis. *Almost Everywhere Convergence II (Proc. Conf. on A.E. Convergence and Probability Theory (Northwestern University, Fall 1989))*. Ed. A. Bellow and R. Jones. Academic Press, New York, 1991, pp. 17–40.

[3] C. Bennett, R. A. DeVore and R. Sharpley. Weak-L^∞ and BMO. *Ann. Math.* **113** (1981), 601–611.

[4] C. Bennett and R. Sharpley. *Interpolation of Operators*. Academic Press, Orlando, FL, 1988.

[5] E. Bishop. *Foundations of Constructive Analysis*. McGraw-Hill, 1967.

[6] J. Bourgain. Pointwise ergodic theorems for arithmetic sets. *Publications Mathematiques, Institut des Hautes Etudes Scientifiques* **69** (1989), 5–45.

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

[8] R. Bradley. Pseudospectral operators and the pointwise ergodic theorem. *Proc. Amer. Math. Soc.* **112** (1991), 863–870.

[9] D. L. Burkholder. Distribution function inequalities for martingales. *Ann. Prob.* **1** (1973), 19–42.

[10] D. L. Burkholder. Martingale transforms. *Ann. Math. Statist.* **37** (1966), 1494–1504.

[11] D. L. Burkholder, B. Davis and R. F. Gundy. Integral inequalities for convex functions of operators on martingales. *Proc. Sixth Berkeley Symp. Math. Stat. Prob.* **2** (1970), 223–240.

[12] A. P. Calderón. Ergodic theory and translation invariant operators. *Proc. Nat. Acad. Sci., USA* **59** (1968), 349–353.

[13] A. P. Calderón. Spaces between L^1 and L^∞ and the theorem of Marcinkiewicz. *Studia Math.* **26** (1966), 273–299.

[14] K. L. Chung. *A Course in Probability Theory*. Harcourt, Brace and World. Inc., New York, 1968.

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press

[15] V. F. Gaposhkin. A theorem on the convergence almost everywhere of a sequence of measurable functions, and its applications to sequences of stochastic integrals. *Math. USSR Sbornik* **33** (1977), 1–17.
[16] V. F. Gaposhkin. Individual ergodic theorem for normal operators on L^2. *Funct. Anal. Appl.* **15** (1981), 14–18.
[17] A. Garsia. *Topics in Almost Everywhere Convergence (Lectures in Advanced Mathematics)*. Markham Publishing, Chicago, 1970.
[18] R. F. Gundy. A decomposition of L^1 bounded martingales. *Ann. Math. Statist.* **39** 134–138.
[19] R. L. Jones. Inequalities for the ergodic maximal function. *Studia Math.* **40** (1977), 111–129.
[20] R. L. Jones, I. Ostrovskii and J. Rosenblatt. Square functions in ergodic theory. *Ergod. Th. & Dynam. Sys.* **16** (1996), 267–305.
[21] S. Kalikow and B. Weiss. Fluctuations of ergodic averages. *Preprint*.
[22] D. Lepingle. La variation d'order p des semi-martingales. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **36** (1976), 295–316.
[23] G. Pisier and Q. Xu. The strong p-variation of martingales and orthogonal series. *Prob. Theory* **77** (1988), 497–514.
[24] N. Riviere. Interpolation a la Marcinkiewicz. *Revista de la Union Mathematica Argentina* **25** (1971), 363–377.
[25] J. Rosenblatt and M. Wierdl. Pointwise ergodic theorems via harmonic analysis. *Proc. Conf. Ergodic Theory (Alexandra, Egypt, 1993) (London Math. Soc. Lecture Notes, 205)*. Cambridge University Press, 1995, pp. 3–151.
[26] E. M. Stein. *Singular Integrals and Differentiablity Properties of Functions*. Princeton University Press, Princeton, NJ, 1970.
[27] H. White. The pointwise ergodic theorem and related analytic inequalities. *Master's Thesis*, University of North Carolina, Chapel Hill, NC, 1989.
[28] A. Zygmund. *Trigonometric Series, Vol. 2*. Second edn. Cambridge University Press, New York, 1977.

https://doi.org/10.1017/S0143385798108349 Published online by Cambridge University Press