

# Square functions in ergodic theory

ROGER L. JONES†

*Department of Mathematics, DePaul University, Chicago, IL 60614, USA*

IOSIF V. OSTROVSKII‡

*Institute for Low Temperature, Physics and Engineering, 47 Lenin Avenue,  
Kharkov 310164, Ukraine*

JOSEPH M. ROSENBLATT‡§

*Department of Mathematics, The Ohio State University, Columbus, OH 43210, USA*

(Received 20 April 1994 and revised 5 January 1995)

*Abstract.* Given the usual averages  $A_n f = \frac{1}{n} \sum_{k=1}^n f \circ \tau^k$  in ergodic theory, let  $n_1 \leq n_2 \leq \dots$  and  $Sf = (\sum_{k=1}^\infty |A_{n_{k+1}} f - A_{n_k} f|^2)^{1/2}$ . There is a strong inequality  $\|Sf\|_2 \leq 25\|f\|_2$  and there is a weak inequality  $m\{Sf > \lambda\} \leq (7000/\lambda)\|f\|_1$ . Related results and questions for other variants of this square function are also discussed.

## 0. Introduction

This article concerns square functions in ergodic theory. However, the methods often concern estimates of Fourier transforms and the behavior of abstract convolution operators. For this reason, many of the results have parallels in real analysis, some of which we describe. In the first section, strong  $L_2$  estimates are the focus. In the second section, weak  $L_1$  estimates are obtained. In the third section, the connection of square functions to maximal functions with random shifts and large deviations is described. We have tried to state what we think are the most interesting unresolved issues connected with this work, as we develop the material.

## 1. Strong $L_2$ estimates for square functions

Let  $(X, \beta, m)$  be a probability space and let  $\tau : X \rightarrow X$  be an invertible  $\beta$ -measurable transformation preserving  $m$ . Given  $f \in L_p = L_p(X, \beta, m)$ , let  $A_n f = \frac{1}{n} \sum_{\ell=1}^n f \circ \tau^\ell$  be the usual average in ergodic theory. The individual ergodic theorem says that there exists  $f^*$  such that  $\lim_{n \rightarrow \infty} A_n f(x) = f^*(x)$  for  $m$  a.e.  $x$ , whenever  $f \in L_p(X)$ ,  $1 \leq p < \infty$ . For this reason, it is obvious that for any increasing sequence  $(n_k)$ , if

† Partially supported by NSF Grant DMS-9302012.

‡ Partially supported by NSF Grant DMS-9103056.

§ Current address: Department of Mathematics, University of Illinois, Urbana, IL 61801, USA.

$f \in L_1$ ,  $\lim_{k \rightarrow \infty} A_{n_{k+1}} f(x) - A_{n_k} f(x) = 0$  a.e. The questions that are addressed here are about ways of discussing the rate at which these differences go to 0; for example, what can be said about  $\sum_{k=1}^{\infty} |A_{n_{k+1}} f(x) - A_{n_k} f(x)|^2$ ?

This same question, but for more general averages than  $A_n f$ , appears in a fundamental way in the work of Bourgain [5, 6] on the convergence of averages along certain subsequences of  $(\tau^\ell : \ell = 1, 2, 3, \dots)$ . For example, the a.e. convergence of

$$\mathcal{A}_n f = \frac{1}{n} \sum_{\ell=1}^n f \circ \tau^{\ell^2}$$

for  $f \in L_2$  was proved by deriving an estimate on the rate of growth of the partial sums

$$\sum_{k=1}^J |A_{n_{k+1}} f - A_{n_k} f|^2.$$

Such questions for sequences other than  $(\ell^2)$  also appeared in Wierdl [21, 22].

These results on square functions suggested the theorem of White [20], see Assani et al [1], that for rapidly growing  $(n_k)$ , the *maximal square function*

$$S^* f = \left( \sum_{k=1}^{\infty} \left( \max_{n_k \leq n \leq n_{k+1}} |A_n f - A_{n_k} f| \right)^2 \right)^{1/2}$$

satisfies a strong  $L_2$ -estimate  $\|S^* f\|_2 \leq C \|f\|_2$  for some constant  $C$ , depending only on  $(n_k)$ . The condition on  $(n_k)$  given there is  $n_{k+1} \geq n_k^8$ . However, by the same argument, one can see that the same result holds for  $n_{k+1} \geq n_k^\alpha$  for some fixed  $\alpha > 1$ . It is not clear if any restriction on  $(n_k)$  is really needed here for this fact to remain true. Indeed, in earlier work, Gaposhkin [12, 13] showed that the same strong  $L_2$ -estimate holds for  $S^*$  if there exists  $\beta \geq \alpha > 1$  such that  $\alpha \leq n_{k+1}/n_k \leq \beta$  for all  $k \geq 1$ . See Bradley [7] for a good exposition of this result in a more general setting.

The first question suggested by such square function bounds is whether there is always a strong  $L_2$ -estimate for the *square function*

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

**THEOREM 1.1.** *For any  $f \in L_2(X)$ ,*

$$\|Sf\|_2 \leq 25 \|f\|_2.$$

This theorem follows immediately from a somewhat more general principle. Let  $U : H \rightarrow H$  be a unitary operator on a Hilbert space. Let  $A_n f = \frac{1}{n} \sum_{\ell=1}^n U^\ell f$  for all  $f \in H$ . For  $f \in H$ , let

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

**THEOREM 1.2.** *For any unitary operator  $U$  and any  $f \in H$ ,  $Sf \leq 25 \|f\|_H$ .*

*Proof of Theorem 1.1.* For any  $N \geq 1$ , the partial sum  $\sum_{k=1}^N |A_{n_{k+1}}f - A_{n_k}f|^2$  is in  $L_1$  if  $f \in L_2$  and

$$\left\| \sum_{k=1}^N |A_{n_{k+1}}f - A_{n_k}f|^2 \right\|_1 = \sum_{k=1}^N \|A_{n_{k+1}}f - A_{n_k}f\|_2^2.$$

Since  $f \mapsto f \circ \tau$  is a unitary operator, Theorem 1.2 says

$$\left\| \sum_{k=1}^N |A_{n_{k+1}}f - A_{n_k}f|^2 \right\|_1 \leq 25^2 \|f\|_2^2.$$

Letting  $N \rightarrow \infty$ , the monotone convergence theorem says

$$\left\| \left( \sum_{k=1}^{\infty} |A_{n_{k+1}}f - A_{n_k}f|^2 \right)^{1/2} \right\|_2^2 = \left\| \sum_{k=1}^{\infty} |A_{n_{k+1}}f - A_{n_k}f|^2 \right\|_1 \leq 25^2 \|f\|_2^2.$$

*Proof of Theorem 1.2.* It suffices to show

$$\sum_{k=1}^N \|A_{n_{k+1}}f - A_{n_k}f\|_H^2 \leq 25^2 \|f\|_H^2$$

for all  $N \geq 1$ . By the spectral theorem for unitary operators, this inequality follows from a similar one on the circle  $T = \{\gamma \in \mathbb{C} : |\gamma| = 1\}$ . Let  $a_n(\gamma) = \frac{1}{n} \sum_{\ell=1}^n \gamma^\ell$ . It suffices to show that for all  $\gamma \in T$ ,

$$\sum_{k=1}^N |a_{n_{k+1}}(\gamma) - a_{n_k}(\gamma)|^2 \leq 25^2.$$

To prove this result, fix  $\gamma \in T$ . Since  $a_n(1) = 1$  for all  $n \geq 1$ , we can assume  $\gamma \neq 1$ . Write  $\gamma = e^{i\theta}$  where  $\theta \in (0, 2\pi)$ . If  $\theta \in [\pi, 2\pi)$ , then  $\bar{\gamma} = e^{i\phi}$  where  $\phi = 2\pi - \theta$  is in  $(0, \pi]$ . Since

$$|a_{n_{k+1}}(\gamma) - a_{n_k}(\gamma)| = |a_{n_{k+1}}(\bar{\gamma}) - a_{n_k}(\bar{\gamma})|,$$

we may assume with no loss of generality that  $\theta \in (0, \pi]$ .

We will split the sum into two sums. First,

$$\begin{aligned} |a_{n_{k+1}}(\gamma) - a_{n_k}(\gamma)|^2 &= \left| \frac{1}{n_{k+1}} \left( \frac{e^{in_{k+1}\theta} - 1}{e^{i\theta} - 1} \right) - \frac{1}{n_k} \left( \frac{e^{in_k\theta} - 1}{e^{i\theta} - 1} \right) \right|^2 \\ &\leq 16 \left| \frac{(e^{in_{k+1}\theta} - 1)}{n_{k+1}\theta} - \frac{(e^{in_k\theta} - 1)}{n_k\theta} \right|^2 \end{aligned}$$

because  $|e^{i\theta} - 1| \geq \frac{1}{4}\theta$  for all  $\theta \in (0, \pi]$ . Then let  $F(z) = (e^{iz} - 1)/z$ . We see that

$$\sum_{k=1}^N |a_{n_{k+1}}(\gamma) - a_{n_k}(\gamma)|^2 \leq 16 \sum_{k=1}^N |F(n_{k+1}\theta) - F(n_k\theta)|^2.$$

Now split this last sum into two sums,  $\Sigma_1$  and  $\Sigma_2$ , where  $\Sigma_1$  is the sum over all  $k = 1, \dots, N$  with  $|n_{k+1}\theta - n_k\theta| < 1$  and  $\Sigma_2$  is the sum over all  $k = 1, \dots, N$  with

$|n_{k+1}\theta - n_k\theta| \geq 1$ . Any term in the first sum estimates by the Cauchy–Schwarz inequality,

$$\begin{aligned} |F(n_{k+1}\theta) - F(n_k\theta)|^2 &= \left| \int_{n_k\theta}^{n_{k+1}\theta} F'(x) dx \right|^2 \\ &\leq |n_{k+1}\theta - n_k\theta| \int_{n_k\theta}^{n_{k+1}\theta} |F'(x)|^2 dx \\ &\leq \int_{n_k\theta}^{n_{k+1}\theta} |F'(x)|^2 dx. \end{aligned}$$

Hence,  $\Sigma_1 \leq \sum_{k=1}^N \int_{n_k\theta}^{n_{k+1}\theta} |F'(x)|^2 dx$ . But  $F'(x) = (ixe^{ix} - (e^{ix} - 1))/x^2$  for  $x > 0$ . Since  $F(z)$  is analytic, and  $|F'(x)| \leq 1/x + 2/x^2$  for all  $x > 0$ , it is clear that  $\int_0^\infty |F'(x)|^2 dx < \infty$ . A straightforward computation gives

$$\Sigma_1 \leq \int_0^\infty |F'(x)|^2 dx \leq 10.$$

To estimate  $\Sigma_2$ , we note that each term in  $\Sigma_2$  is estimated by

$$|F(n_{k+1}\theta) - F(n_k\theta)|^2 \leq 2|F(n_{k+1}\theta)|^2 + 2|F(n_k\theta)|^2.$$

But suppose that  $(n_{k_1}, n_{k_1+1}, \dots, n_{k_L}, n_{k_L+1})$  are the pairs in increasing order appearing in  $\Sigma_2$ . Then  $n_{k_s+1} \geq n_{k_s} + 1/\theta$  for all  $s = 1, \dots, L$ . Because  $n_{k_{s+1}} \geq n_{k_s+1}$ , we have by induction  $n_{k_{s+1}} \geq s/\theta$  for  $s = 1, \dots, L-1$  and  $n_{k_L+1} \geq s/\theta$  for  $s = 1, \dots, L$ . Thus, since  $|F(n\theta)| \leq 1$  always,

$$\begin{aligned} \Sigma_2 &\leq 2|F(n_{k_1}\theta)|^2 + 2 \sum_{s=1}^L |F(n_{k_s+1}\theta)|^2 + 2 \sum_{s=2}^L |F(n_{k_s}\theta)|^2 \\ &\leq 2 + 8 \sum_{s=1}^L \frac{1}{(n_{k_s+1}\theta)^2} + 8 \sum_{s=2}^L \frac{1}{(n_{k_s}\theta)^2} \\ &\leq 2 + \frac{8}{\theta^2} \sum_{s=1}^L \left(\frac{\theta}{s}\right)^2 + \frac{8}{\theta^2} \sum_{s=2}^L \left(\frac{\theta}{s}\right)^2 \\ &\leq 2 + 16 \sum_{s=1}^\infty \frac{1}{s^2} \\ &\leq 29. \end{aligned}$$

Combining the estimates for  $\Sigma_1$  and  $\Sigma_2$  gives

$$\begin{aligned} \sum_{k=1}^N |a_{n_{k+1}}(\gamma) - a_{n_k}(\gamma)|^2 &\leq 16(\Sigma_1 + \Sigma_2) \\ &\leq 16(10 + 29) \leq 25^2. \end{aligned}$$

*Remarks 1.3.* (a) If  $(n_k)$  grows slowly, then the result of Theorem 1.1 is trivial. Indeed, for any  $n_2 \geq n_1$ ,

$$\|A_{n_2}f - A_{n_1}f\|_2 = \left\| \left( \frac{1}{n_2} - \frac{1}{n_1} \right) \sum_{\ell=1}^{n_1} f \circ \tau^\ell + \frac{1}{n_2} \sum_{\ell=n_1+1}^{n_2} f \circ \tau^\ell \right\|_2$$

$$\begin{aligned}
&\leq n_1 \left| \frac{1}{n_2} - \frac{1}{n_1} \right| \|f\|_2 + \frac{n_2 - n_1}{n_2} \|f\|_2 \\
&= 2 \left( \frac{n_2 - n_1}{n_2} \right) \|f\|_2.
\end{aligned}$$

Hence, if  $\rho = \sum_{k=1}^{\infty} (1 - n_k/n_{k+1})^2$ , then  $\|Sf\|_2 \leq 2\sqrt{\rho} \|f\|_2$  for all  $f \in L_2$ . Therefore, if  $(n_k)$  is slowly growing, e.g.  $n_k = k^r$  for some fixed  $r = 1, 2, 3, \dots$ , then  $\rho < \infty$  and the strong inequality of Theorem 1.1 is immediate with  $2\sqrt{\rho}$  in place of 25. Since  $\rho$  can certainly be  $\infty$ , this argument is worthless in general. Actually, it is also the case that  $\rho < \infty$  if and only if  $\sum_{k=1}^{\infty} \sup_{|\gamma|=1} |a_{n_{k+1}}(\gamma) - a_{n_k}(\gamma)|^2$  converges. This explains partly why the estimate in Theorem 1.2 is not generally possible if the terms are estimated uniformly first.

(b) If  $(n_k)$  grows rapidly, then the result of Theorem 1.1 is also well-known because the method of pointwise bounding  $\sum_{k=1}^{\infty} |a_{n_{k+1}}(\gamma) - a_{n_k}(\gamma)|^2$  can be made more explicitly in terms of  $|\gamma - 1|$ . This is the beginning for estimates in Duoandikoetxea and Rubio de Francia [11]. Indeed,  $|a_n(\gamma) - 1| \leq n|\gamma - 1|$  and  $|a_n(\gamma)| \leq 2/n|\gamma - 1|$ . Hence  $|a_{n_{k+1}}(\gamma) - a_{n_k}(\gamma)| \leq 2n_{k+1}|\gamma - 1|$  and  $|a_{n_{k+1}}(\gamma) - a_{n_k}(\gamma)| \leq 4/n_k|\gamma - 1|$ . Thus, for  $L \geq 1$ ,

$$\sum_{k=1}^{\infty} |a_{n_{k+1}}(\gamma) - a_{n_k}(\gamma)|^2 \leq 4|\gamma - 1|^2 \sum_{k=1}^L n_{k+1}^2 + 8 + \frac{16}{|\gamma - 1|^2} \sum_{k=L+3}^{\infty} \frac{1}{n_k^2}$$

because  $\sum_{k=L+1}^{L+2} |a_{n_{k+1}}(\gamma) - a_{n_k}(\gamma)|^2 \leq 8$  for all  $\gamma \in T$ . Thus, we can get a bound as in Theorem 1.1 if there is a constant  $C$  such that for each  $\gamma \in T$ , there is some choice of  $L \geq 1$  with  $|\gamma - 1|^2 \sum_{k=1}^L n_{k+1}^2 \leq C$  and  $(1/|\gamma - 1|^2) \sum_{k=L+3}^{\infty} 1/n_k^2 \leq C$ . For example, if  $(n_k)$  is lacunary with  $\inf_{k \geq 1} n_{k+1}/n_k \geq a > 1$ , then we can choose  $L$  to be the first value with  $n_{L+3} \geq 1/|\gamma - 1|$ . Then

$$|\gamma - 1|^2 \sum_{k=1}^L n_{k+1}^2 \leq |\gamma - 1|^2 C(a) n_{L+2}^2 \leq C(a),$$

where  $C(a) = 1/(a^2 - 1)$ . But then also

$$\frac{1}{|\gamma - 1|^2} \sum_{k=L+3}^{\infty} \frac{1}{n_k^2} \leq \frac{1}{|\gamma - 1|^2} a^2 C(a) \frac{1}{n_{L+3}^2} \leq a^2 C(a).$$

Hence, for such  $(n_k)$ , the strong inequality of Theorem 1.1 holds with a constant  $C_0(a)$  in place of 25. Since  $(n_k)$  may fail to be lacunary, this method may not apply. But moreover, the constant  $C_0(a)$  which is given tends to  $\infty$  as  $a$  tends to 1.

(c) Wittmann pointed out an easy proof of (b) if  $n_k = 2^k$ . First, if  $f \in H$ , then by the parallelogram law

$$\begin{aligned}
\|f - \tfrac{1}{2}(I + T)f\|^2 &= \tfrac{1}{4}\|f - Tf\|^2 \\
&= \tfrac{1}{2}\|f\|^2 + \tfrac{1}{2}\|Tf\|^2 - \tfrac{1}{4}\|f + Tf\|^2 \\
&\leq \|f\|^2 - \|\tfrac{1}{2}(I + T)f\|^2
\end{aligned}$$

for any contraction  $T : H \rightarrow H$ . So

$$\|A_n f - A_{2n} f\|^2 = \|A_n f - \tfrac{1}{2}(I + T^n)A_n f\|^2$$

$$\begin{aligned} &\leq \|A_n f\|^2 - \|\tfrac{1}{2}(I + T^n)A_n f\|^2 \\ &= \|A_n f\|^2 - \|A_{2n} f\|^2. \end{aligned}$$

Therefore,  $\sum_{k=1}^{\infty} \|A_{2^{k+1}} f - A_{2^k} f\|^2 = \|A_2 f\|^2 \leq \|f\|^2$ .

Theorem 1.2 also gives this immediate corollary. Let  $\alpha_n$  be the discrete measure on  $\mathbb{Z}$ ,  $\alpha_n = \frac{1}{n} \sum_{\ell=1}^n \delta_{\ell}$ .

COROLLARY 1.4. *For any  $(n_k)$ ,  $n_k \leq n_{k+1}$  for all  $k \geq 1$ ,*

$$\left\| \left( \sum_{k=1}^{\infty} |(\alpha_{n_{k+1}} - \alpha_{n_k}) * \phi|^2 \right)^{1/2} \right\|_{\ell_2} \leq 25 \|\phi\|_{\ell_2}.$$

This corollary could be used with the Calderón transfer principle to give Theorem 1.1; however, the only proof we know of for Corollary 1.4 would be in the style of the proof of Theorem 1.2, either via the spectral theorem, or by using the Plancherel theorem to recast the estimate on  $T$  (which is essentially the same thing).

Because of the strong parallel between ergodic theorems and differentiation theorems, one suspects there should be an analogous result to Theorem 1.1, but for the Lebesgue derivatives in  $\mathbb{R}$ . Such a result could be obtained by transfer from Corollary 1.4, but it is easier just to repeat the proof. Let  $\varphi_{\varepsilon} = (1/\varepsilon)1_{[0, \varepsilon]} \in L_1(\mathbb{R})$ .

THEOREM 1.5. *For any  $(\varepsilon_k)$ ,  $\varepsilon_k \geq \varepsilon_{k+1} > 0$  for all  $k \geq 1$ ,*

$$\left\| \left( \sum_{k=1}^{\infty} |(\varphi_{\varepsilon_{k+1}} - \varphi_{\varepsilon_k}) * f|^2 \right)^{1/2} \right\|_2 \leq 7 \|f\|_2$$

for all  $f \in L_2(\mathbb{R})$ .

*Proof.* It suffices to show that for all  $x \geq 0$ ,

$$\sum_{k=1}^{\infty} |\hat{\varphi}_{\varepsilon_{k+1}}(x) - \hat{\varphi}_{\varepsilon_k}(x)|^2 \leq 49.$$

Here

$$\hat{\varphi}_{\varepsilon}(x) = \frac{1}{\varepsilon} \left( \frac{1 - e^{-ix\varepsilon}}{ix} \right).$$

Hence,

$$\begin{aligned} |\hat{\varphi}_{\varepsilon_{k+1}}(x) - \hat{\varphi}_{\varepsilon_k}(x)| &= \left| \frac{(1 - e^{-ix\varepsilon_{k+1}})}{x\varepsilon_{k+1}} - \frac{(1 - e^{-ix\varepsilon_k})}{x\varepsilon_k} \right| \\ &= |F(x\varepsilon_{k+1}) - F(x\varepsilon_k)| \end{aligned}$$

where  $F(r) = (1 - e^{-ir})/r$  for all  $r > 0$ , and  $F(0) = 0$ . Now split this sum  $\sum_{k=1}^{\infty} |\hat{\varphi}_{\varepsilon_{k+1}}(x) - \hat{\varphi}_{\varepsilon_k}(x)|^2$  into two sums,  $\Sigma_1$  and  $\Sigma_2$ , where  $\Sigma_1$  is over  $k$  such that  $|(\varepsilon_{k+1} - \varepsilon_k)x| < 1$  and  $\Sigma_2$  is over  $k$  such that  $|(\varepsilon_{k+1} - \varepsilon_k)x| \geq 1$ . As before,

$$\begin{aligned} \Sigma_1 &\leq \sum_{k=1}^{\infty} \int_{\varepsilon_{k+1}x}^{\varepsilon_kx} |F'(r)|^2 dr \\ &\leq \int_0^{\infty} |F'(r)|^2 dr \\ &\leq 10. \end{aligned}$$

To estimate  $\Sigma_2$ , let  $(k_s : s = 1, 2, 3, \dots, L)$  be the indices in increasing order with  $(\varepsilon_{k_s} - \varepsilon_{k_{s+1}})x \geq 1$ . Then  $\varepsilon_{k_{L-s}} \geq (s+1)/x$  for  $s = 0, \dots, L-1$  and  $\varepsilon_{k_{L-s}+1} \geq s/x$  for  $s = 1, \dots, L-1$ . Hence,

$$\begin{aligned} \Sigma_2 &\leq 2 \sum_{s=1}^L |F(\varepsilon_{k_s+1}x)|^2 + 2 \sum_{s=1}^L |F(\varepsilon_{k_s}x)|^2 \\ &\leq 2|F(\varepsilon_{k_L+1}x)|^2 + 8 \sum_{s=1}^{L-1} \frac{1}{(\varepsilon_{k_s+1}x)^2} + 8 \sum_{s=1}^L \frac{1}{(\varepsilon_{k_s}x)^2}. \end{aligned}$$

Since  $|F(r)| \leq 1$  for all  $r > 0$ ,

$$\begin{aligned} \Sigma_2 &\leq 2 + 8 \sum_{s=1}^{L-1} \frac{1}{x^2} \frac{x^2}{(L-s)^2} + 8 \sum_{s=1}^L \frac{1}{x^2} \frac{x^2}{(L-s+1)^2} \\ &\leq 2 + 16 \sum_{s=1}^{\infty} \frac{1}{s^2} \\ &\leq 29. \end{aligned}$$

Combining the estimates for  $\Sigma_1$  and  $\Sigma_2$  gives

$$\sum_{k=1}^{\infty} |\hat{\varphi}_{\varepsilon_{k+1}}(x) - \hat{\varphi}_{\varepsilon_k}(x)|^2 \leq 10 + 29 = 39.$$

*Remark 1.6.* (a) The reversal in direction of summation that occurs in the estimate of  $\Sigma_2$  in the proof of Theorem 1.5, compared with the proof of Theorem 1.2, is typical in estimating expressions related to Lebesgue differentiation, as opposed to similar ones in ergodic theory. If there is a weight accompanying the index, this causes the parallel of these two analyses to break down. See Rosenblatt and Wierdl [16] where large deviation theorems are proved in ergodic theory, which generally fail for Lebesgue derivatives.

(b) Theorem 1.5 can be generalized to other sequences that form an approximate identity. For example, assume  $dm = p(x) dx$  where  $p \in L_1(\mathbb{R})$ ,  $p \geq 0$ ,  $\int p(x) dx = 1$ . Assume that for some  $s > 0$ ,  $p(x) \exp(s|x|)$  is in  $L_2(\mathbb{R})$ . Let  $m_a$  be the dilation  $m_a(E) = m((1/a)E)$ . Then for any  $W = \{(a_k, b_k) : k \geq 1\}$  which are pairwise-disjoint,

$$\left\| \left( \sum_{k=1}^{\infty} |(m_{b_k} - m_{a_k}) * f|^2 \right)^{1/2} \right\|_2 \leq \sqrt{\frac{51}{s}} \|p(x) \exp(s|x|)\|_{2,x} \|f\|_2$$

for all  $f \in L_2(\mathbb{R})$ . This applies, of course, to any  $p$  which has bounded support. The proof of this theorem, and related ones for square functions similar to this, will appear in another article.

(c) Theorem 1.5 suggests the following natural question. Let  $E_k f$  denote the conditional expectation for  $f \in L_1(\mathbb{R})$  given by

$$E_k f(x) = \frac{1}{(1/2^k)} \int_{j/2^k}^{(j+1)/2^k} f(t) dt$$

whenever  $x \in [j/2^k, j + 1/2^k)$ ,  $j \in \mathbb{Z}$ . It is well-known that  $\lim_{k \rightarrow \infty} E_k f(x) = f(x)$  a.e. Indeed, this is the martingale convergence theorem for the martingale  $(E_k)$ . By Burkholder's inequality [8],

$$\left\| \left( \sum_{k=1}^{\infty} |E_{k+1} f - E_k f|^2 \right)^{1/2} \right\|_2 \leq C \|f\|_2,$$

for all  $f \in L_2(\mathbb{R})$ . Comparison of this fact with Theorem 1.5, and the close analogy of  $E_k f$  with  $\varphi_{1/2^k} * f$ , suggest the question whether  $\sum_{k=1}^{\infty} |E_k f - \varphi_{1/2^k} * f|^2 < \infty$  a.e., for all  $f \in L_2(\mathbb{R})$ ? This would be an obvious consequence of a strong inequality

$$\left\| \left( \sum_{k=1}^{\infty} |E_k f - \varphi_{1/2^k} * f|^2 \right)^{1/2} \right\|_2 \leq C \|f\|_2.$$

Is there such an inequality? In a similar fashion, certain reversed martingales in  $\ell_1(\mathbb{Z})$  dominate the usually averaged  $\alpha_n * f = \frac{1}{n} \sum_{\ell=1}^n f(j - \ell)$  for  $f \in \ell_1(\mathbb{Z})$ ; see Rosenblatt and Wierdl [16]. Is there a strong inequality for the square function of the differences between  $\alpha_{2^n} * f$  and the associated reversed martingale on  $\ell_2(\mathbb{Z})$ , similar to the one suggested above? Recently, we have seen that the answer to these questions is affirmative; these results will appear elsewhere in joint work of Jones, Kaufman, Rosenblatt and Wierdl.

There are two main directions of generalization of Theorem 1.1 of interest: one is to other operators and other averages, the other is to square functions of block maxima as in Bourgain [6]. First, it is straightforward to improve Theorem 1.1 to a general contraction.

**THEOREM 1.7.** *Let  $T$  be a contraction on a Hilbert space  $H$ . Let  $(n_k)$  be a sequence in  $\mathbb{Z}^+$  with  $n_k \leq n_{k+1}$  for all  $k \geq 1$ . Let  $A_n(T)f = \frac{1}{n} \sum_{\ell=1}^n T^\ell f$  for all  $f \in H$ . Then*

$$\left( \sum_{k=1}^{\infty} \|A_{n_{k+1}}(T)f - A_{n_k}(T)f\|_H^2 \right)^{1/2} \leq 25 \|f\|_H$$

for all  $f \in H$ .

*Proof.* By the dilation theorem, see Sz-Nagy and Foias [19], there exists a Hilbert space  $L$  containing  $H$  as a closed subspace, an orthogonal projection  $P : L \rightarrow H$ , and a unitary mapping  $U : L \rightarrow L$  with  $P U^\ell f = T^\ell f$  for all  $\ell \geq 0$  and  $f \in H$ . But then for  $f \in H$ ,

$$\begin{aligned} \sum_{k=1}^N \|A_{n_{k+1}}(T)f - A_{n_k}(T)f\|_H^2 &= \sum_{k=1}^N \|P(A_{n_{k+1}}(U) - A_{n_k}(U)f)\|_H^2 \\ &= \|P\|^2 \sum_{k=1}^N \|A_{n_{k+1}}(U)f - A_{n_k}(U)f\|_H^2 \\ &\leq 25^2 \|f\|_H^2 \end{aligned}$$

by Theorem 1.2 and the fact that  $\|P\| \leq 1$ . □



Square functions for other averages are also of interest. For example, recently the behavior of iterates of an average of the form  $\mu f(x) = \sum_{\ell=-\infty}^{\infty} \mu(\ell) f(\tau^\ell x)$ , with respect to a probability measure  $\mu$  on  $\mathbb{Z}$ , have received considerable attention in ergodic theory. The associated square functions is the following: fix  $(n_k)$ ,  $n_k \leq n_{k+1}$ , and let  $S_\mu f = (\sum_{k=1}^{\infty} |\mu^{n_{k+1}} f - \mu^{n_k} f|^2)^{1/2}$ . The problem is to obtain a strong inequality with suitable conditions on  $(n_k)$  and  $\text{spec}(\mu) = \text{cl}\{\hat{\mu}(\gamma) : \gamma \in T\}$ . Here is an example.

**THEOREM 1.8.** *A necessary and sufficient condition for there to be a constant  $C$  such that*

$$\left\| \left( \sum_{k=1}^{\infty} |\mu^{n_{k+1}} f - \mu^{n_k} f|^2 \right)^{1/2} \right\|_2 \leq C \|f\|_2$$

*for all  $f \in L_2$ , and all dynamical systems  $(X, \beta, m, \tau)$ , is that there is a closed circular disc  $C_\rho$  of radius  $1 - \rho$  centered at  $\rho > 0$  in  $\mathbb{C}$  with  $\text{spec}(\mu) \subset C_\rho$ .*

*Proof.* First, using the spectral theorem to obtain the strong inequality, it suffices to have  $\sum_{k=1}^{\infty} |\hat{\mu}^{n_{k+1}}(\gamma) - \hat{\mu}^{n_k}(\gamma)|^2 \leq C$  for all  $\gamma \in T$ . That is,  $|\hat{\mu}(\gamma) - 1|^2 \sum_{k=1}^{\infty} |\hat{\mu}(\gamma)|^{2k}$  must be bounded. If  $|\hat{\mu}(\gamma)| < 1$  for all  $\gamma \neq 1$ , then this is the same as having

$$\frac{|\hat{\mu}(\gamma) - 1|^2}{1 - |\hat{\mu}(\gamma)|^2} |\hat{\mu}(\gamma)|^2 \leq C$$

for all  $\gamma \in T$ ,  $\gamma \neq 1$ . But if  $\text{spec}(\mu) \subset C_\rho$ ,  $C_\rho$  a circular disc as above, then for some constant  $K_\rho$ ,  $\sup_{\gamma \in T, \gamma \neq 1} |\hat{\mu}(\gamma) - 1|^2 / (1 - |\hat{\mu}(\gamma)|) \leq K_\rho$ . So the condition on  $\text{spec}(\hat{\mu})$  is sufficient for the strong inequality. Conversely, if there is a strong inequality, valid for all dynamical systems, then  $|\hat{\mu}(\gamma) - 1|^2 \sum_{k=1}^{\infty} |\hat{\mu}(\gamma)|^{2k}$  is uniformly bounded. Hence,  $|\hat{\mu}(\gamma)| < 1$  except for  $\gamma = 1$ , and  $|\hat{\mu}(\gamma) - 1|^2 |\hat{\mu}(\gamma)|^2 \leq C(1 - |\hat{\mu}(\gamma)|^2)$  for all  $\gamma \in T$ . But then, for some  $\rho > 0$ ,  $\text{spec}(\mu) \subset C_\rho$ .  $\square$

**Remark 1.9.** If for some aperiodic (e.g. ergodic) non-atomic finite dynamical system there is a strong inequality as in Theorem 1.8, then there is such an inequality for all dynamical systems with the same constant. This can be seen by the Conze principle. See Bellow *et al* [2] or Rosenblatt and Wierdl [16] for a discussion of this principle and examples of its use. Alternatively, one can use the Rokhlin Lemma and the Calderón transfer principle to prove the same thing.

It turns out that the spectral criterion of Theorem 1.8 is implied by the more familiar one of strict aperiodicity. We say that a probability measure  $\mu$  on  $\mathbb{Z}$  is *strictly aperiodic* if  $|\hat{\mu}(\gamma)| < 1$  for all  $\gamma \in T$ ,  $\gamma \neq 1$ , i.e.,  $\mu$  is strictly aperiodic if and only if its support is not contained in a proper arithmetic progression on  $\mathbb{Z}$ .

Let

$$I[\hat{\mu}(e^{it})] = \frac{|\hat{\mu}(e^{it}) - 1|^2}{1 - |\hat{\mu}(e^{it})|^2} \quad \text{and} \quad J[\hat{\mu}(e^{it})] = \frac{(\text{Im } \hat{\mu}(e^{it}))^2}{1 - \text{Re } \hat{\mu}(e^{it})}.$$

**THEOREM 1.10.** *If  $\mu$  is a strictly aperiodic probability measure on  $\mathbb{Z}$ , then  $I[\hat{\mu}(e^{it})]$  is bounded.*

To prove this, we first prove two lemmas.

LEMMA 1.11. *If there exists a  $k \in \mathbb{Z}$  such that  $I[\hat{\mu}(e^{it})e^{ikt}]$  is bounded, then  $I[\hat{\mu}(e^{it})]$  is bounded.*

*Proof.* We have

$$\begin{aligned} I[\hat{\mu}(e^{it})] &= \frac{|(\hat{\mu}(e^{it})e^{ikt} - 1)e^{-ikt} + (e^{-ikt} - 1)|^2}{1 - |\hat{\mu}(e^{it})e^{ikt}|^2} \\ &\leq 2I[\hat{\mu}(e^{it})e^{ikt}] + 2\frac{|e^{-ikt} - 1|^2}{1 - |\hat{\mu}(e^{it})|^2}. \end{aligned}$$

Denoting by  $\mu_1$  the measure defined by the equality  $\hat{\mu}_1 = |\hat{\mu}|^2$ , we have

$$\begin{aligned} \liminf_{t \rightarrow 0} \frac{1 - |\hat{\mu}(e^{it})|^2}{|e^{-ikt} - 1|^2} &= \frac{1}{k^2} \liminf_{t \rightarrow 0} \frac{1 - \operatorname{Re} \hat{\mu}_1(e^{it})}{t^2} \\ &= \frac{1}{k^2} \liminf_{t \rightarrow 0} \sum_{n \in \mathbb{Z}} \mu_1(n) \frac{1 - \cos nt}{t^2} \\ &\geq \frac{1}{k^2} \liminf_{t \rightarrow 0} \sum_{n=-N}^N \mu_1(n) \frac{1 - \cos nt}{t^2} \\ &= \frac{1}{2k^2} \sum_{n=-N}^N \mu_1(n) n^2 \end{aligned}$$

for any natural  $N$ . If  $N$  is large enough, the last term is strictly positive. That proves the lemma.  $\square$

LEMMA 1.12. *If  $\mu$  is strictly aperiodic, then  $I[\hat{\mu}(e^{it})]$  is bounded over  $t \neq 0$  if and only if*

$$j[\mu] = \limsup_{t \rightarrow 0} J[\hat{\mu}(e^{it})] < 2.$$

*Proof.* We have

$$I[\hat{\mu}(e^{it})] = \frac{1 - \operatorname{Re} \hat{\mu}(e^{it}) + J[\hat{\mu}(e^{it})]}{1 + \operatorname{Re} \hat{\mu}(e^{it}) - J[\hat{\mu}(e^{it})]},$$

hence

$$\limsup_{t \rightarrow 0} I[\hat{\mu}(e^{it})] = \frac{j[\mu]}{2 - j[\mu]}.$$

Because  $\mu$  is strictly aperiodic, this proves the lemma.  $\square$

*Proof of Theorem 1.10.* By Lemma 1.11 we may suppose that  $\mu(0) > 0$  without loss of generality. By Lemma 1.12 we may restrict ourselves to proving  $j[\mu]$  is bounded away from 2. Consider the identity

$$\begin{aligned} &\sum_{n \in \mathbb{Z}} \mu(n) \sin^2 \frac{nt}{2} \sum_{n \in \mathbb{Z}} \mu(n) \cos^2 \frac{nt}{2} - \left( \sum_{n \in \mathbb{Z}} \mu(n) \sin \frac{nt}{2} \cos \frac{nt}{2} \right)^2 \\ &= \frac{1}{2} \sum_{n, m \in \mathbb{Z}} \mu(n) \mu(m) \left( \sin \frac{(n-m)t}{2} \right)^2. \end{aligned}$$

Neglecting all terms with  $m \neq 0$  on the right-hand side, we see that the right-hand side is not less than

$$\frac{1}{2}\mu(0) \sum_{n \in \mathbb{Z}} \mu(n) \sin^2 \frac{nt}{2}.$$

Note that

$$1 - \operatorname{Re} \hat{\mu}(e^{it}) = 2 \sum_{n \in \mathbb{Z}} \mu(n) \sin^2 \frac{nt}{2},$$

$$\operatorname{Im} \hat{\mu}(e^{it}) = 2 \sum_{n \in \mathbb{Z}} \mu(n) \sin \frac{nt}{2} \cos \frac{nt}{2}.$$

Therefore we have

$$\frac{1}{2}(1 - \operatorname{Re} \hat{\mu}(e^{it})) \sum_{n \in \mathbb{Z}} \mu(n) \cos^2 \frac{nt}{2} - (\frac{1}{2} \operatorname{Im} \hat{\mu}(e^{it}))^2 \geq \frac{1}{2}\mu(0)\frac{1}{2}(1 - \operatorname{Re} \hat{\mu}(e^{it})),$$

hence

$$2(1 - \operatorname{Re} \hat{\mu}(e^{it})) - (\operatorname{Im} \hat{\mu}(e^{it}))^2 \geq \mu(0)(1 - \operatorname{Re} \hat{\mu}(e^{it})),$$

and

$$J[\hat{\mu}(e^{it})] \leq 2 - \mu(0).$$

So, Theorem 1.10 is proved.  $\square$

**COROLLARY 1.13.** *If  $\mu$  is a strictly aperiodic probability measure on  $\mathbb{Z}$ , then for some constant  $C$ ,*

$$\left\| \left( \sum_{k=1}^{\infty} |\mu^{k+1} f - \mu^k f|^2 \right)^{1/2} \right\|_2 \leq C \|f\|_2$$

for all  $f \in L_2$ .

*Proof.* The estimate in Theorem 1.10 is precisely what is needed for the spectral hypothesis in Theorem 1.8 to hold.  $\square$

**Remark 1.14.** The bound on  $\hat{\mu}$  which is inherent in Corollary 1.13 is exactly what is needed to give the ideal improvement of the subsequence theorem in Gaposhkin and Rosenblatt [14]; no moment condition on  $\mu$  is really needed for the subsequence results. For example, if  $\mu$  is just a strictly aperiodic probability measure on  $\mathbb{Z}$  and  $\tau$  is invertible, then for any subsequence  $(n_m)$  with  $n_{m+1} \geq n_m^\alpha$  for some fixed  $\alpha > 1$ , the averages  $\mu^{n_m} f(x)$  converge a.e. for all  $f \in L_2(X)$ . See [14] for the details and why this improved estimate for the spectrum of  $\mu$  gives such a subsequence theorem.

Corollary 1.13 concerns  $S_\mu f$  at one extreme, where  $(n_k)$  grows slowly. At another extreme, with  $(n_k)$  arbitrary, we have this result. For  $\alpha < \pi$ , let  $\mathcal{S}_\alpha$  be the usual Stolz region for non-tangential convergence at 1, with aperture  $\alpha$ , that is,  $\mathcal{S}_\alpha$  can be characterized as a region on which there is a bound  $|1 - z|/(1 - |z|) \leq C_\alpha$  for all  $z \in \mathcal{S}_\alpha$ .

**THEOREM 1.15.** *If  $\mu$  is a probability measure on  $\mathbb{Z}$  and  $\operatorname{spec}(\mu) \subset \mathcal{S}_\alpha$ , then there is a constant  $C_\alpha$  such that for all  $(n_k)$ ,  $n_k \leq n_{k+1}$  for  $k \geq 1$ , we have*

$$\left\| \left( \sum_{k=1}^{\infty} |\mu^{n_{k+1}} f - \mu^{n_k} f|^2 \right)^{1/2} \right\|_2 \leq C_\alpha \|f\|_2$$

for all  $f \in L_2$ , and for all dynamical systems  $(X, \beta, m, \tau)$ .

*Proof.* If  $0 \leq r \leq 1$ , then

$$\sum_{k=1}^{\infty} (r^{n_{k+1}} - r^{n_k})^2 \leq 4 \sum_{k=1}^{\infty} (r^{n_k} - r^{n_{k+1}}) = 4r^{n_1} \leq 4.$$

But then if  $z \in \mathcal{S}_\alpha$ , we have

$$\begin{aligned} \sum_{k=1}^{\infty} |z^{n_{k+1}} - z^{n_k}|^2 &= \sum_{k=1}^{\infty} |z|^{2n_k} |z^{n_{k+1}-n_k} - 1|^2 \\ &\leq \sum_{k=1}^{\infty} |z|^{2n_k} |z - 1|^2 (1 + \dots + |z|^{n_{k+1}-n_k-1})^2 \\ &\leq C_\alpha^2 \sum_{k=1}^{\infty} |z|^{2n_k} (1 - |z|)^2 (1 + \dots + |z|^{n_{k+1}-n_k-1})^2 \\ &= C_\alpha^2 \sum_{k=1}^{\infty} |z|^{2n_k} (1 - |z|^{n_{k+1}-n_k})^2 \\ &= C_\alpha^2 \sum_{k=1}^{\infty} (|z|^{n_k} - |z|^{n_{k+1}})^2 \\ &\leq 4C_\alpha^2 \end{aligned}$$

by letting  $r = |z|$  and using the estimate above.

But now by the spectral theorem, to prove the strong inequality above, it suffices to prove  $\sum_{k=1}^{\infty} |\hat{\mu}(\gamma)^{n_{k+1}} - \hat{\mu}(\gamma)^{n_k}|^2 \leq C_\alpha^2$ . But  $\text{spec}(\mu) \subset \mathcal{S}_\alpha$  and the estimate above gives such a result.  $\square$

*Remark 1.16.* It is probably the case that the only way the strong inequality of Theorem 1.15 can hold for some (all) dynamical system(s) is to have  $\text{spec}(\mu)$  in some Stolz angle. See Bellow *et al* [3] for other facts about  $\mu$  which has a spectrum that is restricted as in Theorem 1.15.

Another version of this problem is to fix the probability measure  $\mu$ , and depending on  $\text{spec}(\mu)$ , obtain conditions on  $(n_k)$  for which the strong inequality holds. Corollary 1.13 shows that  $n_k = k$  will do for any strictly aperiodic measure. Actually, if  $(n_k)$  is more rapidly growing, then it will also work.

**THEOREM 1.17.** *Let  $(n_k)$  be a sequence of natural numbers such that  $n_{k+1} \geq n_k^p$  for some  $p > 1$ . Then, for any strictly aperiodic probability measure  $\mu$ , the sum*

$$S(\gamma) = \sum_{k=1}^{\infty} |\hat{\mu}^{n_{k+1}}(\gamma) - \hat{\mu}^{n_k}(\gamma)|^2$$

*is uniformly bounded on the unit circle  $|\gamma| = 1$ .*

*Proof.* This argument is similar to the one in Remark 1.3b. Define for any natural  $L$ ,

$$S_L = \sum_{k=1}^L |\hat{\mu}^{n_{k+1}} - \hat{\mu}^{n_k}|^2 = \sum_{k=1}^L |\hat{\mu}|^{2n_k} |1 - \hat{\mu}|^2 |1 + \dots + \hat{\mu}^{n_{k+1}-n_k-1}|^2$$

$$\begin{aligned} &\leq |1 - \hat{\mu}|^2 \sum_{k=1}^L (n_{k+1} - n_k)^2 \leq |1 - \hat{\mu}|^2 \left( \sum_{k=1}^L n_{k+1} - n_k \right)^2 \\ &\leq C|1 - \hat{\mu}|^2 n_{L+1}^2. \end{aligned}$$

Moreover, using estimates as in the proof of Theorem 1.15, we have

$$\begin{aligned} R_L = \sum_{k=L+3}^{\infty} |\hat{\mu}^{n_{k+1}} - \hat{\mu}^{n_k}|^2 &\leq \frac{|1 - \hat{\mu}|^2}{(1 - |\hat{\mu}|)^2} \sum_{k=L+3}^{\infty} (|\hat{\mu}|^{n_k} - |\hat{\mu}|^{n_{k+1}})^2 \\ &\leq \frac{4|1 - \hat{\mu}|^2}{(1 - |\hat{\mu}|)^2} |\hat{\mu}|^{n_{L+3}}. \end{aligned}$$

Since for every strictly aperiodic  $\mu$  we have by Theorem 1.10 that

$$|1 - \hat{\mu}|^2 \leq C(1 - |\hat{\mu}|),$$

then

$$\begin{aligned} S_L &\leq C(1 - |\hat{\mu}|)n_{L+1}^2 \\ R_L &\leq 4C(1 - |\hat{\mu}|)^{-1}|\hat{\mu}|^{n_{L+3}}. \end{aligned}$$

Choose  $L$  in the following way:

$$n_{L+1} \leq \frac{1}{(1 - |\hat{\mu}|)^{1/2}} < n_{L+2}.$$

Then  $S_L \leq C$ . Since  $p > 1$ , we can choose  $d \geq 1$  with  $p^d > 2$ . Then  $n_{L+2+d} \geq n_{L+2}^{p^d}$ . But, in addition,

$$\begin{aligned} \log \frac{|\hat{\mu}|^{n_{L+2+d}}}{1 - |\hat{\mu}|} &= -n_{L+2+d} \log \frac{1}{|\hat{\mu}|} + \log \frac{1}{1 - |\hat{\mu}|} \\ &\leq -n_{L+2}^{p^d} \log \frac{1}{|\hat{\mu}|} + \log \frac{1}{1 - |\hat{\mu}|} \\ &\leq -\frac{1}{(1 - |\hat{\mu}|)^{p^d/2}} \log \frac{1}{|\hat{\mu}|} + \log \frac{1}{1 - |\hat{\mu}|} \\ &\leq -\frac{1}{2}(1 - |\hat{\mu}|)^{1-p^d/2} + \log \frac{1}{1 - |\hat{\mu}|} \rightarrow -\infty \quad \text{as } |\hat{\mu}| \rightarrow 1. \end{aligned}$$

Thus,  $R_{L+d-1}$  is bounded. But the choice of  $d$  is independent of  $L$  and so  $S \leq S_L + 4(d+1) + R_{L+d-1}$  is also bounded.  $\square$

Despite Theorem 1.17 and Corollary 1.13, not every sequence  $(n_k)$  will do for every strictly aperiodic measure  $\mu$ .

*Example 1.18.* Let  $\mu = \frac{1}{2}(\delta_0 + \delta_1)$ . Then

$$\hat{\mu}(e^{it}) = \frac{1}{2}(1 + e^{-it}) = \frac{1}{2}e^{-it/2}(e^{it/2} + e^{-it/2}).$$

So  $\hat{\mu}(e^{it}) = e^{-it/2} \cos(t/2)$ . Let  $S(e^{it}) = \sum_{k=1}^{\infty} |\hat{\mu}^{n_{k+1}}(e^{it}) - \hat{\mu}^{n_k}(e^{it})|^2$ . Choose  $n_k = 1 + 3 + \dots + 3^{k-1} = (3^k - 1)/2$  and  $t_p = 2\pi/3^p$ . Then

$$S(e^{it_p}) = \sum_{k=1}^{\infty} \left( \cos \frac{t_p}{2} \right)^{3^k-1} \left( 1 + \left( \cos \frac{t_p}{2} \right)^{2 \cdot 3^k} - 2 \left( \cos \frac{t_p}{2} \right)^{3^k} \cos \left( 3^k \frac{t_p}{2} \right) \right)$$

$$\begin{aligned}
&\geq \sum_{k=p}^{\infty} \left( \cos \frac{t_p}{2} \right)^{3^k} \left( 1 + \left( \cos \frac{t_p}{2} \right)^{2 \cdot 3^k} + 2 \left( \cos \frac{t_p}{2} \right)^{3^k} \right) \\
&\geq \sum_{k=p}^{2p} \left( \cos \frac{t_p}{2} \right)^{3^k} \geq p \left( \cos \frac{t_p}{2} \right)^{3^{2p}}
\end{aligned}$$

The last expression tends to  $+\infty$  when  $p$  tends to  $+\infty$  since

$$\lim_{p \rightarrow \infty} \left( \cos \frac{t_p}{2} \right)^{3^{2p}} = \exp(-\pi^2/2).$$

*Remark 1.19.* (a) It is not hard to compute examples that link the choice of  $(n_k)$  to the shape of  $\text{spec}(\mu)$ . The computations in Bellow *et al* which lead to [2, Theorem 1.14] actually give a prescription for such examples. However, it would be better to resolve what is really the general pattern. For example, let  $G(z) = \sum_{k=1}^{\infty} z^{n_k}$ ,  $|z| \leq 1$ . In terms of the mapping properties of  $G$ , can we determine precise conditions on  $\text{spec}(\mu)$  which are necessary and sufficient for Theorem 1.17 with that choice of  $(n_k)$  and  $\mu$ ?

(b) In a similar manner to the proof of Theorem 1.7, one can show that if  $\mu(T)f = \sum_{\ell=-\infty}^{\infty} \mu(\ell)T^{\ell}f$  (with  $T$  invertible if  $\text{supp } \mu \not\subset \mathbb{Z}^+$ ), then for the choice of  $(n_k)$  in Theorem 1.17, and for any contraction  $T$  on a Hilbert space  $H$ ,

$$\sum_{k=1}^{\infty} \|\mu(T)^{n_{k+1}}f - \mu(T)^{n_k}f\|_H^2 \leq C^2 \|f\|_H$$

for all  $f \in H$ .

The question of getting strong estimates for square functions of block maxima is also quite worthwhile, especially because it has the potential of giving stronger inequalities than the usual maximal inequalities in the individual ergodic theorem. Fix  $(n_k)$ ,  $n_k \leq n_{k+1}$ . The square maximal function in question is as before:

$$S^*f = \left( \sum_{k=1}^{\infty} \left( \max_{n_k \leq n \leq n_{k+1}} |A_n f - A_{n_k} f| \right)^2 \right)^{1/2}.$$

We will also want to discuss, in the next section, a somewhat more restricted version of  $S^*f$ . Let  $M$  be a sequence  $(m_k : k \geq 1)$  in  $\mathbb{Z}^+$ , then

$$S_M^*f = \left( \sum_{k=1}^{\infty} \left( \max_{\substack{n_k \leq n \leq n_{k+1} \\ n \in M}} |A_n f - A_{n_k} f| \right)^2 \right)^{1/2}$$

In Assani *et al* [1], a theorem of White's is proved, which has some precedents in Bourgain [5, 6].

**THEOREM 1.20.** *Let  $(n_k)$  satisfy  $n_{k+1} \geq n_k^{\alpha}$  for some  $\alpha > 1$ . Then there is a constant  $C = C(\alpha) < \infty$  such that  $\|S^*f\|_2 \leq C(\alpha)\|f\|_2$  for all  $f \in L_2$ , and for all dynamical systems.*

*Remarks 1.21.* (a) The actual hypothesis in [1] is that  $n_{k+1} \geq n_k^8$ . However, by the same proof (or by passing to subsequences of  $(n_k)$ ), the result holds for any  $\alpha > 1$ . Because

of Theorem 1.1, it is not unreasonable to hope that Theorem 1.20 remains true for all  $(n_k)$ .

(b) The same proof as given in [1] shows that for  $(\varepsilon_k)$ , if for some  $\alpha > 1$ ,  $\varepsilon_{k+1} \leq \varepsilon_k^\alpha$  for all  $k \geq 1$ , then for all  $f \in L_2(\mathbb{R})$ ,

$$\left\| \left( \sum_{k=1}^{\infty} \sup_{\varepsilon_{k+1} \leq \varepsilon \leq \varepsilon_k} |(\varphi_\varepsilon - \varphi_{\varepsilon_k}) * f|^2 \right)^{1/2} \right\|_2 \leq C \|f\|_2.$$

Here  $C$  depends only on  $\alpha$ .

The same type of result on block maxima is at least true with no restriction on  $(n_k)$ , if one uses  $S_M^* f$  instead of  $Sf$ . See Theorem 4.10 in Rosenblatt and Wierdl [17].

**THEOREM 1.22.** *Let  $(n_k)$  be any increasing sequence and let  $M$  be a lacunary sequence. Then there is a constant  $C < \infty$ , depending only on the degree of lacunarity of  $M$ , such that  $\|S_M^* f\|_2 \leq C(\alpha) \|f\|_2$  for all  $f \in L_2$ , and for all dynamical systems.*

## 2. Weak $L_1$ estimates for square functions

In this section, two different approaches to obtaining weak inequalities in  $L_1$  for the square function will be given. The first approach only applies to lacunary  $(n_k)$ , but is also better for obtaining strong  $L_p$  estimates and will be used for other purposes in §3. The second approach uses the Calderón–Zygmund decomposition. Both approaches require having a strong inequality somewhere at the outset.

In Jones [15], it is shown that for

$$Sf = \left( \sum_{k=1}^{\infty} |(A_{k+1} - A_k)f|^2 \right)^{1/2}$$

there is a weak estimate,  $m\{Sf > \lambda\} \leq (C/\lambda) \|f\|_1$ , valid for some constant  $C < \infty$  and arbitrary  $f \in L_1$ . The same method can be tried in general, but only seems to yield a strong  $L_p$  inequality for  $1 < p < \infty$ , and that only when  $(n_k)$  is a polynomial function of  $k$ . This is one reason for the interest in the following result.

**THEOREM 2.1.** *Suppose  $(n_k)$  is lacunary, with  $n_{k+1}/n_k \geq \beta > 1$  for all  $k \geq 1$ . Then there is a constant  $C(\beta)$  such that for all  $f \in L_1$ ,*

$$m \left\{ \left( \sum_{k=1}^{\infty} |(A_{n_{k+1}} - A_{n_k})f|^2 \right)^{1/2} > \lambda \right\} \leq \frac{C(\beta)}{\lambda} \|f\|_1.$$

*Proof.* We use a theorem on vector-valued Calderón–Zygmund operators from Benedek *et al* [4]. This result says that we can get a weak  $L_1$  inequality from a strong  $L_2$  inequality, and certain properties of the operator in question. See also Rubio de Francia *et al* [18].

First, by the Calderón transfer principle, it suffices to prove the analogous result in  $\mathbb{Z}$ , namely, with  $\alpha_n = \frac{1}{n} \sum_{\ell=1}^n \delta_\ell$ ,

$$\# \left\{ s : \left( \sum_{k=1}^{\infty} |(\alpha_{n_{k+1}} - \alpha_{n_k}) * \varphi(s)|^2 \right)^{1/2} > \lambda \right\} \leq \frac{C(\beta)}{\lambda} \|\varphi\|_{\ell_1(\mathbb{Z})}$$

for all  $\varphi \in \ell_1(\mathbb{Z})$ . (See Bellow *et al* [2] or Rosenblatt and Wierdl [16] for some general forms of the Calderón transfer principle [10] which would work here.) However, it is equivalent to show that if  $\varphi_n = \frac{1}{n} 1_{[0,n]}$ , that with respect to the Lebesgue measure on  $\mathbb{R}$ , if  $f \in L_1(\mathbb{R})$ , then

$$m \left\{ \left( \sum_{k=1}^{\infty} |(\varphi_{n_{k+1}} - \varphi_{n_k}) * f|^2 \right)^{1/2} > \lambda \right\} \leq \frac{C(\beta)}{\lambda} \|f\|_1.$$

(See Bellow *et al* [2] where a similar transfer from  $\mathbb{R}$  to  $\mathbb{Z}$  is used to translate a theorem of Duoandikoetxea and Rudio de Francia [11] from  $\mathbb{R}$  to  $\mathbb{Z}$ .)

Now define the kernel operator  $K : \mathbb{R} \rightarrow \ell_2(\mathbb{Z}^+)$  by

$$K(x) = \left( \frac{1}{n_{k+1}} 1_{[0, n_{k+1}]}(x) - \frac{1}{n_k} 1_{[0, n_k]}(x) : k = 1, 2, 3, \dots \right).$$

This is the appropriate operation in this case to which to apply the main result from Benedek *et al* [4]. Indeed,  $Af = \int K(x-y)f(y)dy$  has

$$\|Af\|_{\ell_2(\mathbb{Z}^+)} = \left( \sum_{k=1}^{\infty} |(\varphi_{n_{k+1}} - \varphi_{n_k}) * f|^2 \right)^{1/2}$$

and so an estimate on  $m\{\|Af\|_{\ell_2(\mathbb{Z}^+)} > \lambda\}$  is exactly what is required.

Theorem 1.1 and the definition of  $A$  show that the proof of Theorems 1 and 2 in [4] give Theorem 2.1 here, if  $K$  satisfies the Hörmander condition:

$$\int_{|x| > 4|y|} \|K(x-y) - K(x)\|_{\ell_2(\mathbb{Z}^+)} dx \leq C_2$$

where  $C_2 < \infty$  is independent of  $y \in \mathbb{R}$ .

To check the Hörmander condition in this case, we need to evaluate  $|\varphi_{n_k}(x-y) - \varphi_{n_k}(x)|$  for  $|x| > 4|y|$ . Let us first take the case  $x > 4y$ ,  $y > 0$ . Then

$$\begin{aligned} |\varphi_{n_k}(x-y) - \varphi_{n_k}(x)| &= \frac{1}{n_k} |1_{[y, y+n_k]}(x) - 1_{[0, n_k]}(x)| \\ &= \begin{cases} 0 & y > n_k \\ \frac{1}{n_k} (-1_{[0, y]}(x) + 1_{[n_k, y+n_k]}(x)) & y \leq n_k \end{cases}, \end{aligned}$$

because if  $x > 4y$ , then  $x > n_k$  and  $x > y + n_k$  when  $y > n_k$ . So for  $x > 4y$ ,

$$|\varphi_{n_k}(x-y) - \varphi_{n_k}(x)| = \frac{1}{n_k} 1_{[n_k, y+n_k]}(x)$$

if  $n_k \geq y$ , and it is 0 otherwise. This means that for fixed  $y$ ,

$$\begin{aligned} \int_{x > 4y} \|K(x-y) - K(x)\|_{\ell_2(\mathbb{Z}^+)} dx &\leq \int_{x > 4y} \left( \sum_{k=1}^{\infty} |\varphi_{n_{k+1}}(x-y) - \varphi_{n_{k+1}}(x)|^2 \right)^{1/2} dx \\ &\quad + \int_{x > 4y} \left( \sum_{k=1}^{\infty} |\varphi_{n_k}(x-y) - \varphi_{n_k}(x)|^2 \right)^{1/2} dx \\ &= \int_{x > 4y} \left( \sum_{y \leq n_{k+1}} \frac{1}{n_{k+1}^2} 1_{[n_{k+1}, n_{k+1}+y]}(x) \right)^{1/2} dx \end{aligned}$$



$$\begin{aligned}
 & + \int_{x>4y} \left( \sum_{y \leq n_k} \frac{1}{n_k^2} 1_{[n_k, n_k+y]}(x) \right)^{1/2} dx \\
 & \leq 2 \int \left( \sum_{y \leq n_k} \frac{1}{n_k^2} 1_{[n_k, n_k+y]}(x) \right)^{1/2} dx \\
 & \leq 2 \int \sum_{y \leq n_k} \frac{1}{n_k} 1_{[n_k, n_k+y]}(x) dx \\
 & = 2y \sum_{y \leq n_k} \frac{1}{n_k}.
 \end{aligned}$$

But since  $n_{k+1}/n_k \geq \beta > 1$ , there is a constant  $C(\beta)$  such that  $\sum_{y \leq n_k} 1/n_k \leq C(\beta)/y$ . Hence,  $\int_{x>4y} \|K(x-y) - K(x)\|_{\ell_2(\mathbb{Z}^+)} dx \leq 2C(\beta)$  for all  $y > 0$ .

Similar calculations can be used in the other cases. For instance, if  $y \leq 0$  and  $x > 4|y|$ , then we need to compute

$$|\varphi_{n_k}(x-y) - \varphi_{n_k}(x)| = \frac{1}{n_k} |1_{[y, y+n_k]}(x) - 1_{[0, n_k]}(x)|$$

again. But for similar reasons as before, this is  $(1/n_k)1_{[n_k+y, n_k]}(x)$  for  $x > 4|y|$ . Hence

$$\int_{x>4|y|} \|K(x-y) - K(x)\|_{\ell_2(\mathbb{Z}^+)} dx \leq 2C(\beta)$$

for all  $y$ .

Finally, for  $y > 0$  and  $x < 0$ ,  $\varphi_{n_k}(x-y) - \varphi_{n_k}(x) = 0$ . Also for  $y < 0$  and  $x < 0$ ,

$$\varphi_{n_k}(x-y) - \varphi_{n_k}(x) = \frac{1}{n_k} 1_{[y, y+n_k]}(x).$$

But if also  $|x| > 4|y|$ , then  $x < 4y$  and so  $1_{[y, y+n_k]}(x) = 0$  again. That is, if  $x < 0$  and  $|x| > 4|y|$ , then  $\|K(x-y) - K(x)\|_{\ell_2(\mathbb{Z}^+)} = 0$  for any  $y$ .

The conclusion is that for  $(n_k)$  lacunary, the Hörmander condition holds. Hence, for  $(n_k)$  lacunary, the associated square function is weak  $L_1$ .  $\square$

*Remarks 2.2.* (a) The condition needed for Hörmander's inequality in the proof is really

$$\int_{|x|>4(y)} \left( \sum_{n_k \geq |y|} \frac{1}{n_k^2} 1_{[n_k, n_k+y]}(x) \right)^{1/2} dx \leq C.$$

This is only true if  $(n_k)$  is essentially lacunary (a finite union of lacunary sequences) because it implies that if  $n(y) = \#\{n_k : n_k \leq y\}$ , then  $n(2y) - n(y)$  is bounded.

(b) The question is whether Theorem 2.1 holds without any condition on  $(n_k)$ . For example, if  $n_k = k^2$ , then

$$\begin{aligned}
 Sf &= \left( \sum_{k=1}^{\infty} \left| \frac{-(2k+1)}{(k+1)^2 k^2} \sum_{\ell=1}^{k^2} f \circ \tau^\ell + \frac{1}{(k+1)^2} \sum_{\ell=k^2+1}^{(k+1)^2} f \circ \tau^\ell \right|^2 \right)^{1/2} \\
 &\leq C \left( \sum_{k=1}^{\infty} \frac{1}{k^2} |A_{k^2} f|^2 \right)^{1/2} + C \left( \sum_{\ell=1}^{\infty} \frac{1}{k^2} |A_{2k+1} f \circ \tau^{k^2}|^2 \right)^{1/2}
 \end{aligned}$$

since  $\sum_{k=1}^{\infty} 1/k^2 < \infty$ ; the first term is dominated by  $C \sup_{k \geq 1} |A_k f|$ . Because  $\sup_{k \geq 1} |A_k f|$  is weak  $L_1$ , it is easy to see that with  $n_k = k^2$ , the square function of Theorem 2.1 is weak  $L_1$  if and only if  $(\sum_{k=1}^{\infty} (1/k^2) |A_k f \circ \tau^{k^2}|^2)^{1/2}$  is weak  $L_1$ . This is very interesting because the method in Jones [15] does not apply here. Also,  $(A_k f \circ \tau^{k^2} : k \geq 1)$  does not converge a.e., so  $\sup_k |A_k f \circ \tau^{k^2}|$  is not weak  $L_1$ . But in Rosenblatt and Wierdl [16] it is shown that

$$\sum_{k=1}^{\infty} m\{A_k f > \lambda k\} \leq \frac{C}{\lambda} \|f\|_1.$$

Hence,  $\lim_{k \rightarrow \infty} A_k f \circ \tau^{k^2}/k = 0$  a.e., for  $f \in L_1$ . The unresolved question is whether  $(A_n f \circ \tau^{k^2}/k)$  goes to 0 fast enough for  $\sum_{k=1}^{\infty} (A_k f \circ \tau^{k^2}/k)^2 < \infty$  a.e.? See Theorem 2.6 for a proof that this is indeed true.

The method in Duoandikoetxea and Rubio de Francia [11] shows that the square function  $Sf$  of Theorem 2.1 is strong  $L_p$  for all  $p$ ,  $1 < p < \infty$ . However, the weak inequality does not follow from their method directly. The method of Theorem 2.1 also gives this strong  $L_p$  result.

**THEOREM 2.3.** *If  $(n_k)$  is lacunary, then there is a constant  $C$  such that*

$$\left\| \left( \sum_{k=1}^{\infty} |(A_{n_{k+1}} - A_{n_k}) f|^2 \right)^{1/2} \right\|_p \leq C \|f\|_p$$

for all  $f \in L_p$ ,  $1 < p < \infty$ , and all dynamical systems.

*Proof.* See Duoandikoetxea and Rubio de Francia [11] or the proof of Theorem 2 in Benedek et al [4].  $\square$

It would be quite worthwhile to also apply the method of Theorem 2.1 to the maximal square function  $S^* f$ . Unfortunately, this does not seem to work. Instead, the best that can be obtained by this method, in a straightforward manner, is this more restricted version which applies to  $S_M^* f$  for suitable  $M$ .

**THEOREM 2.4.** *Suppose  $(n_k)$  and  $M = (m_k)$  are lacunary. Then  $S_M^* f$  is weak  $L_1$  and strong  $L_p$  for  $1 < p < \infty$ , for all dynamical systems.*

*Proof.* As in Theorem 2.1, by Theorem 1.22, it suffices to show that a certain Banach space valued convolution operator  $K$  satisfies the Hörmander condition. In this case, the operator  $K$  is given from  $B_1 = \mathbb{R}$  to  $B_2$ , an  $\ell_2$  sum of finite-dimensional  $\ell_\infty$  spaces. Specifically, we write the general element  $d \in B_2$  as  $d = ((d_m : n_k \leq m \leq n_{k+1}, m \in M) : k > 1)$ ; then

$$\|d\|_{B_2} = \left( \sum_{k=1}^{\infty} \left( \max_{\substack{n_k \leq m \leq n_{k+1} \\ m \in M}} |d_m| \right)^2 \right)^{1/2}$$

Then let  $K : B_1 \rightarrow B_2$  given by

$$K(x) = \left( \left( \frac{1}{m} 1_{[0, m]} - \frac{1}{n_k} 1_{[0, n_k]} : n_k \leq m \leq n_{k+1}, m \in M \right) : k \geq 1 \right).$$

The condition that is needed to prove Theorem 2.4 then becomes

$$\int_{|x| \geq 4|y|} \|K(x-y) - K(x)\|_{B_2} dx \leq C_2 \quad \text{for all } y \in \mathbb{R}.$$

Also,  $M$  is lacunary and  $(n_k)$  is lacunary. Therefore, as in the proof of Theorem 2.1, there is a constant  $C_2 = C(\alpha, \beta)$ , where  $\alpha$  is the lacunarity constant for  $(n_k)$  and  $\beta$  is the lacunarity constant for  $M$ , such that the above Hörmander condition holds.  $\square$

*Remark 2.5.* (a) It is clear from the manner in which the constant  $C(\alpha, \beta)$  is determined that  $C(\alpha, \beta)$  only becomes unbounded as  $\alpha \downarrow 1$  and/or  $\beta \downarrow 1$ . So there is some  $C < \infty$  such that  $C \geq C(\alpha, \beta)$  whenever  $\alpha \geq 2$  and  $\beta \geq 2$ . It follows that the weak inequality of Theorem 2.4 directly gives the usual weak inequality in the ergodic theorem. That is, we fix  $n_1 = 1$  and  $n_2 \geq 1$ . Then by Theorem 2.4,

$$m \left\{ \max_{n_1 \leq 2^k \leq n_2} |(A_{2^k} - A_{n_1})f| > \lambda \right\} \leq \frac{C}{\lambda} \|f\|_1,$$

where  $C$  does not depend on the choice of  $n_2$ . Hence,

$$m \left\{ \sup_{k \geq 1} |A_{2^k} f| > \lambda \right\} \leq \frac{2C + 2}{\lambda} \|f\|_1$$

by applying the monotone convergence theorem. Of course, for any  $n \geq 1$ , if  $2^k \leq n \leq 2^{k+1}$ , then  $|A_n f| \leq 2A_{2^{k+1}}|f|$ . So

$$m \left\{ \sup_{n \geq 1} |A_n f| > \lambda \right\} \leq \frac{4C}{\lambda} \|f\|_1,$$

for all  $\lambda > 0$  and  $f \in L_1$ . However, this is no advantage because (1) the derivation of Theorem 2.4 is a long way around to get the usual weak  $L_1$  inequality of the ergodic theorem, and (2) maximal inequalities from the ergodic theorem are used twice in the proof, once in White's theorem [20] and essentially once (in the form of the Hardy–Littlewood maximal inequality) in the derivation of the Calderón–Zygmund decomposition in the proof of Theorem 2 in [4].

(b) It would be really striking to obtain directly in  $(X, \beta, m)$  or  $\mathbb{Z}$ , a weak inequality for  $S^*f$  which was valid with a constant independent of the choice of  $(n_k)$ . However, Theorem 2.4 is probably the correct ergodic theoretical version of this corresponding result for martingales. Using Burkholder [8] and Burkholder *et al* [9], Burkholder has commented in a private communication that there exists a constant such that for all martingales  $(f_n)$  which are conditional expectations  $E(f | \beta_n)$  for some  $f \in L_1$ , for any  $(n_k)$ ,

$$m \left\{ \left( \sum_{k=1}^{\infty} \left( \max_{n_k \leq n \leq n_{k+1}} |f_n - f_{n_k}| \right)^2 \right)^{1/2} > \lambda \right\} \leq \frac{C}{\lambda} \|f\|_1.$$

(c) There is an analogous result for Lebesgue differentiation to Theorem 2.4. The proof uses the same method, based on the inequality in Remark 1.21(b).

(d) It would be worthwhile to extend Theorem 2.4 to other square functions, in the same way that Theorem 1.1 was extended. For instance, it is not clear for which

probability measures  $\mu$  on  $\mathbb{Z}$  and  $(n_k)$ , the square functions

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

is weak  $L_1$ . Also, it would be worthwhile to know if

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

can be weak  $L_1$ . For such measures  $\mu$ , Theorem 1.10 in Bellow *et al* [3] gives a very simple proof that  $\sup_{n \geq 1} |\mu^n f|$  is weak  $L_1$ .

We now consider the same question of a weak inequality on  $L_1$ , but we use, instead of the previous singular integral method, the Calderón–Zygmund decomposition directly.

**THEOREM 2.6.** *Let  $(n_k)$  denote any increasing sequence of positive integers. Let*

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

*Then  $Sf$  is weak type  $(1, 1)$  and strong type  $(p, p)$  for  $1 < p \leq 2$ .*

The proof will follow from a number of lemmas. We use both  $|B|$  and  $\#B$  to denote the cardinality of a set.

**LEMMA 2.7.** (The Calderón–Zygmund decomposition.) *Let  $f$  be a function in  $\ell_1(\mathbb{Z})$ . Let  $\lambda > 0$ . Then we can write  $f = g + b$  where  $g \in \ell^2$ , and*

$$\text{CZ} - 1 \quad \|g\|_{\ell_1} \leq \|f\|_{\ell_1},$$

$$\text{CZ} - 2 \quad \|g\|_{\infty} \leq 2\lambda,$$

$$\text{CZ} - 3 \quad b = \sum_i b_i(x) \text{ where each } b_i \text{ satisfies :}$$

$$\text{CZ} - 3(a) \quad b_i \text{ is supported on an interval } B_i.$$

$$\text{CZ} - 3(b) \quad \sum_j b_i(j) = 0 \text{ for each } i.$$

$$\text{CZ} - 3(c) \quad \frac{1}{|B_i|} \sum_{j \in B_i} |b_i(j)| \leq 4\lambda \text{ and } \lambda \leq \frac{1}{|B_i|} \sum_{j \in B_i} |f(j)|$$

$$\text{CZ} - 3(d) \quad B_i \cap B_j = \emptyset \text{ for each } i \neq j.$$

*Remark.* Note that the above imply

$$\sum_i |B_i| \leq \frac{1}{\lambda} \sum_i \|b_i\|_{\ell^1} \leq \frac{1}{\lambda} \|f\|_{\ell_1}.$$

Also, if  $\lambda \geq \|f\|_{\infty}$ , then we take  $f = g$  and  $b = 0$ .

*Proof.* Find an interval  $I$  of length  $2^n$  with  $n$  so large that  $\frac{1}{|I|} \sum_{j \in I} |f(j)| \leq \lambda$  and  $|f(j)| \leq \lambda$  for  $j \notin I$ . Now divide  $I$  into two equal pieces,  $I_1$  and  $I_2$ . Look at the

average of  $|f|$  over each of these pieces. If the average is more than  $\lambda$  on any interval, keep that interval. If the average is less than  $\lambda$  then divide that interval into two equal intervals, and repeat the procedure. The procedure ends with a collection of intervals, the average over which is at least  $\lambda$ , but because of the construction, the average is no more than  $2\lambda$ . Off the selected intervals, the function is at most  $\lambda$ , since clearly any point where  $|f|$  is more than  $\lambda$  would be in some selected interval, possibly an interval containing only the point itself. Denote the selected intervals by  $B_1, B_2, \dots$ . Now define  $g(j) = f(j)$  for  $j$  not in any of the selected intervals. For  $j$  in a selected interval  $B_i$ , define

$$g(j) = \frac{1}{|B_i|} \sum_{r \in B_i} f(r).$$

Define  $b(j) = f(j) - g(j)$ . From the construction each of the required properties follow easily, with  $b_i = b1_{B_i}$ .  $\square$

Let  $\tilde{B}_i$  denote an interval of length  $5|B_i|$  and with the same center as  $B_i$ . Let  $\tilde{B} = \cup_i \tilde{B}_i$ . Let  $j \notin \tilde{B}$ . We have

$$\begin{aligned} Sb(j)^2 &= \sum_{k=1}^{\infty} |A_{n_{k+1}}b(j) - A_{n_k}b(j)|^2 \\ &= \sum_{k=1}^{\infty} \left| A_{n_{k+1}} \left( \sum_i b_i(j) \right) - A_{n_k} \left( \sum_i b_i(j) \right) \right|^2 \\ &= \sum_{k=1}^{\infty} \left| \sum_i (A_{n_{k+1}}b_i(j) - A_{n_k}b_i(j)) \right|^2. \end{aligned}$$

Note that for any  $i$  for which the average includes all the points in  $B_i$ , the average is 0 by CZ-3(b) above. Thus for each fixed  $k$ ,  $A_{n_{k+1}}b_i(j) - A_{n_k}b_i(j)$  is non-zero only if  $j+1 \in B_i$ , i.e. at least the average starts in  $B_i$ , or one of  $j+n_k \in B_i$ , or  $j+n_{k+1} \in B_i$ , i.e. at least one of the averages ends in  $B_i$ . The first possibility, starting in  $B_i$ , is excluded since  $j \notin \tilde{B}$ . Hence for each fixed  $k$  and  $j$ ,  $A_{n_{k+1}}b_i(j) - A_{n_k}b_i(j) \neq 0$  for at most 2 values of  $i$ , an ending value for  $A_{n_{k+1}}b_i(j)$  and an ending value for  $A_{n_k}b_i(j)$ . Thus we know

$$\begin{aligned} Sb(j)^2 &\leq 2 \sum_{k=1}^{\infty} \sum_i |A_{n_{k+1}}b_i(j) - A_{n_k}b_i(j)|^2 \\ &= 2 \sum_i Sb_i(j)^2. \end{aligned}$$

We now have

$$\begin{aligned} \#\{j \mid Sb(j) > \lambda\} &= \#\{j \mid Sb(j)^2 > \lambda^2\} \\ &= \#\{j \mid j \notin \tilde{B}, Sb(j)^2 > \lambda^2\} + \#\{j \mid j \in \tilde{B}, Sb(j)^2 > \lambda^2\} \\ &\leq \frac{1}{\lambda^2} \sum_{j \notin \tilde{B}} Sb(j)^2 + |\tilde{B}|. \end{aligned}$$

We have

$$\frac{1}{\lambda^2} \sum_{j \notin \tilde{B}} Sb(j)^2 \leq \frac{1}{\lambda^2} \sum_{j \notin \tilde{B}} 2 \sum_i Sb_i(j)^2 \leq 2 \sum_i \frac{1}{\lambda^2} \sum_{j \notin \tilde{B}_i} Sb_i(j)^2.$$

For this reason, the following lemma is needed.

LEMMA 2.8. *For each  $i$  we have*

$$\frac{1}{\lambda^2} \sum_{j \notin \tilde{B}_i} S b_i(j)^2 \leq 64 |B_i|.$$

*Proof.* Because translation by an integer is measure preserving, we can assume, without loss of generality, that  $B_i = [0, N - 1]$  where  $|B_i| = N$ . Note that since we only need to consider  $j \in \tilde{B}_i^c$ , we do not need to consider  $j \in (-2N, 3N)$ , and since we are only looking at forward averages, we only need to consider  $j \in (-\infty, -2N]$ . To have a non-zero value of  $A_{n_{k+1}} b_i(j)$  we must reach the support of  $b_i$ . Hence, we must have  $n_{k+1} + j \geq 0$ . Thus,  $n_{k+1} \geq |j|$ . But we might have  $n_k + j \geq 0$  or  $n_k + j < 0$  for that particular value of  $k$ . Let  $n(j)$  be the smallest integer such that  $n_{n(j)+1} \geq |j|$ . Then  $n_{n(j)} + j < 0$  and so we have arranged  $b_i(j + r) = 0$  for all  $r = 1, \dots, n_{n(j)}$ .

$$\begin{aligned} S b_i(j)^2 &= \sum_{n_{k+1} \geq |j|} |A_{n_{k+1}} b_i(j) - A_{n_k} b_i(j)|^2 \leq \left( \sum_{n_{k+1} \geq |j|} |A_{n_{k+1}} b_i(j) - A_{n_k} b_i(j)| \right)^2 \\ &\leq \left( \sum_{n_{k+1} \geq |j|} \left\{ \left| \left( \frac{1}{n_{k+1}} - \frac{1}{n_k} \right) \sum_{r=1}^{n_k} b_i(j+r) \right| + \frac{1}{n_{k+1}} \sum_{r=n_k+1}^{n_{k+1}} |b_i(j+r)| \right\} \right)^2 \\ &\leq \left( \sum_{n_{k+1} \geq |j|} \left( \frac{1}{n_k} - \frac{1}{n_{k+1}} \right) \sum_{r=1}^{n_k} |b_i(j+r)| + \sum_{n_{k+1} \geq |j|} \frac{1}{n_{k+1}} \sum_{r=n_k+1}^{n_{k+1}} |b_i(j+r)| \right)^2 \\ &\leq 2 \left( \sum_{n_{k+1} \geq |j|} \left( \frac{1}{n_k} - \frac{1}{n_{k+1}} \right) \sum_{r=1}^{n_k} |b_i(j+r)| \right)^2 \\ &\quad + 2 \left( \sum_{n_{k+1} \geq |j|} \frac{1}{n_{k+1}} \sum_{r=n_k+1}^{n_{k+1}} |b_i(j+r)| \right)^2 \\ &\leq 2 \left( \sum_{k=n(j)+1}^{\infty} \left( \frac{1}{n_k} - \frac{1}{n_{k+1}} \right) \sum_{r=0}^{N-1} |b_i(r)| \right)^2 \\ &\quad + 2 \left( \sum_{k=n(j)}^{\infty} \frac{1}{n_{k+1}} \sum_{r=n_k+1}^{n_{k+1}} |b_i(j+r)| \right)^2 \\ &\leq 2 \left( \frac{1}{n_{n(j)+1}} N \frac{1}{N} \sum_{r=0}^{N-1} |b_i(r)| \right)^2 + 2 \left( \sum_{k=n(j)}^{\infty} \frac{1}{n_{n(j)+1}} \sum_{r=n_k+1}^{n_{k+1}} |b_i(j+r)| \right)^2 \\ &\leq 32 \left( \frac{1}{n_{n(j)+1}} N \lambda \right)^2 + 2 \left( \frac{1}{n_{n(j)+1}} N \frac{1}{N} \sum_{r=0}^{N-1} |b_i(r)| \right)^2 \\ &\leq 32 \left( \frac{1}{n_{n(j)+1}} N \lambda \right)^2 + 32 \left( \frac{1}{n_{n(j)+1}} N \lambda \right)^2 \\ &\leq 64 \left( \frac{1}{n_{n(j)+1}} N \lambda \right)^2. \end{aligned}$$

We now consider

$$\begin{aligned}
 \sum_{j \notin \tilde{B}_i} S b_i(j)^2 &\leq \sum_{j \leq -2N} 64 \left( \frac{1}{n_n(j)+1} N \lambda \right)^2 \\
 &\leq \sum_{j \leq -2N} 64 \left( \frac{1}{j} N \lambda \right)^2 \\
 &\leq 64 N^2 \lambda^2 \sum_{j \leq -2N} \frac{1}{j^2} \\
 &\leq 64 \lambda^2 N \\
 &= 64 \lambda^2 |B_i|.
 \end{aligned}$$

□

*Proof of Theorem 2.6.* We first establish the weak type  $(1, 1)$  inequality. We have

$$\#\{j : Sf > \lambda\} \leq \#\left\{j : Sg > \frac{\lambda}{2}\right\} + \#\left\{j : Sb > \frac{\lambda}{2}\right\}.$$

For the first term we have, by Theorem 1.1,

$$\begin{aligned}
 \#\left\{j : Sg > \frac{\lambda}{2}\right\} &\leq \frac{4}{\lambda^2} \sum_j Sg(j)^2 \\
 &\leq \frac{2500}{\lambda^2} \sum_j g(j)^2 \\
 &\leq \frac{5000}{\lambda^2} \sum_j \lambda |g(j)| \\
 &\leq \frac{5000}{\lambda} \sum_j |f(j)|.
 \end{aligned}$$

For the second term we have

$$\#\left\{j : Sb(j) > \frac{\lambda}{2}\right\} = \#\left\{j : Sb(j)^2 > \left(\frac{\lambda}{2}\right)^2\right\} \leq \frac{4}{\lambda^2} \sum_{j \notin \tilde{B}} Sb(j)^2 + |\tilde{B}|.$$

We use Lemma 2.8 to conclude that the first term is dominated by  $512 \sum_i |B_i|$ . The second term in this expression is controlled by the same type of sum. Thus

$$\#\left\{j : Sb(j) > \frac{\lambda}{2}\right\} \leq 516 \sum_i |B_i| \leq \frac{1032}{\lambda} \|f\|_{\ell^1}.$$

Hence,

$$\#\{j : Sf > \lambda\} \leq \frac{6032}{\lambda} \|f\|_{\ell^1}.$$

The transfer principle of Calderón gives the theorem with the same constant. The fact that  $S$  is strong type  $(p, p)$ ,  $1 < p \leq 2$ , now follows by interpolation between the weak type  $(1, 1)$  just established and the strong type  $(2, 2)$  of Theorem 1.1. □

*Remark 2.9.* It is not yet clear whether  $Sf$  is also always going to satisfy a strong  $L_p$  estimate for  $2 < p < \infty$ .

The same argument as in Theorem 2.6 will give weak inequalities for other square functions, if there is a strong inequality in  $L_2$ .

**THEOREM 2.10.** *Let  $\{n_k\}$  denote an increasing sequence of integers and define*

$$S^*f(x) = \left( \sum_{k=1}^{\infty} \sup_{n_k \leq n \leq n_{k+1}} |A_n f(x) - A_{n_k} f(x)|^2 \right)^{1/2}$$

*If there is a constant  $C$  such that  $\|S^*f\|_2 \leq C\|f\|_2$  for all  $f \in L_2(X)$  then  $S^*$  is weak type  $(1, 1)$ .*

*Proof.* The proof will follow as in Theorem 2.6. Write  $f = g + b$  as before, and use the hypothesis that  $\|S^*g\|_2 \leq C\|g\|_2$  to handle  $g$ . Thus it remains to control  $S^*b$ .

We first need to show that for  $j \notin \tilde{B}$  we have  $S^*b(j)^2 \leq 2 \sum_i S^*b_i(j)$ . Fix  $j \notin \tilde{B}$ . Since  $b$  is supported in a finite interval, for each  $j$  and  $k$  there is an integer  $n(j, k, b) \in [n_k, n_{k+1}]$  such that

$$\sum_{k=1}^{\infty} \sup_{n_k \leq n \leq n_{k+1}} |A_n b(j) - A_{n_k} b(j)|^2 = \sum_{k=1}^{\infty} |A_{n(j,k,b)} b(j) - A_{n_k} b(j)|^2.$$

Using this fact, we argue as before:

$$\begin{aligned} S^*b(j)^2 &= \sum_{k=1}^{\infty} |A_{n(j,k,b)} b(j) - A_{n_k} b(j)|^2 \\ &= \sum_{k=1}^{\infty} \left| A_{n(j,k,b)} \left( \sum_i b_i(j) \right) - A_{n_k} \left( \sum_i b_i(j) \right) \right|^2 \\ &= \sum_{k=1}^{\infty} \left| \sum_i (A_{n(j,k,b)} b_i(j) - A_{n_k} b_i(j)) \right|^2. \end{aligned}$$

As before,  $A_{n(j,k,b)} b_i(j) - A_{n_k} b_i(j)$  can be non-zero for at most two values of  $i$ , if  $j + n(j, k, b) \in B_i$  and if  $j + n_k \in B_i$ . Thus

$$S^*b(j)^2 \leq 2 \sum_{k=1}^{\infty} \sum_i |A_{n(j,k,b)} b_i(j) - A_{n_k} b_i(j)|^2 = 2 \sum_i S^*b_i(j)^2.$$

We now need the analog of Lemma 2.8. This follows easily once we understand the proof of Lemma 2.8. The only real change is to replace  $n_{k+1}$  with  $n(j, k, b)$  in several places, and use the fact that  $n_k < n(j, k, b) \leq n_{k+1}$ . Also, here  $n_{n(j)} < |j|$  and  $n_{n(j)+1} \geq |j|$ , but also  $n(j, k, b) \geq |j|$ , if the term being considered is not zero.

$$\begin{aligned} S^*b_i(j)^2 &= \sum_{n_{k+1} \geq |j|} |A_{n(j,k,b)} b_i(j) - A_{n_k} b_i(j)|^2 \\ &\leq \left( \sum_{n_{k+1} \geq |j|} |A_{n(j,k,b)} b_i(j) - A_{n_k} b_i(j)| \right)^2 \end{aligned}$$



$$\begin{aligned}
 &\leq \left( \sum_{n_{k+1} \geq |j|} \left\{ \left| \left( \frac{1}{n(j, k, b)} - \frac{1}{n_k} \right) \sum_{r=1}^{n_k} b_i(j+r) \right| \right. \right. \\
 &\quad \left. \left. + \frac{1}{n(j, k, b)} \sum_{r=n_k+1}^{n(j, k, b)} |b_i(j+r)| \right\}^2 \right) \\
 &\leq \left( \sum_{n_{k+1} \geq |j|} \left( \frac{1}{n_k} - \frac{1}{n(j, k, b)} \right) \sum_{r=1}^{n_k} |b_i(j+r)| \right. \\
 &\quad \left. + \sum_{n_{k+1} \geq |j|} \frac{1}{n(j, k, b)} \sum_{r=n_k+1}^{n(j, k, b)} |b_i(j+r)| \right)^2 \\
 &\leq 2 \left( \sum_{n_{k+1} \geq |j|} \left( \frac{1}{n_k} - \frac{1}{n(j, k, b)} \right) \sum_{r=1}^{n_k} |b_i(j+r)| \right)^2 \\
 &\quad + 2 \left( \sum_{n_{k+1} \geq |j|} \frac{1}{n(j, k, b)} \sum_{r=n_k+1}^{n(j, k, b)} |b_i(j+r)| \right)^2 \\
 &\leq 2 \left( \sum_{k=n(j)+1}^{\infty} \left( \frac{1}{n_k} - \frac{1}{n(j, k, b)} \right) \sum_{r=1}^N |b_i(r)| \right)^2 \\
 &\quad + 2 \left( \sum_{k=n(j)}^{\infty} \frac{1}{n(j, k, b)} \sum_{r=n_k+1}^{n(j, k, b)} |b_i(j+r)| \right)^2 \\
 &\leq 2 \left( \frac{1}{n_{n(j)+1}} N \frac{1}{N} \sum_{r=0}^{N-1} |b_i(r)| \right)^2 + 2 \left( \sum_{k=n(j)}^{\infty} \frac{1}{|j|} \sum_{r=n_k+1}^{n_{k+1}} |b_i(j+r)| \right)^2 \\
 &\leq 32 \left( \frac{1}{n_{n(j)+1}} N \lambda \right)^2 + 2 \left( \frac{1}{|j|} N \frac{1}{N} \sum_{r=0}^{N-1} |b_i(r)| \right)^2 \\
 &\leq 32 \left( \frac{1}{n_{n(j)+1}} N \lambda \right)^2 + 32 \left( \frac{1}{|j|} N \lambda \right)^2 \\
 &\leq 64 \left( \frac{1}{|j|} N \lambda \right)^2.
 \end{aligned}$$

The rest of the proof is the same as in the proof of Theorem 2.6. □

This result combines with White's theorem and Gaposhkin's theorem to give:

**THEOREM 2.11.** *For any  $(n_k)$  with  $n_{k+1} > n_k^\rho$  for some  $\rho > 1$ , or  $\beta \geq n_{k+1}/n_k \geq \alpha$  for some  $\beta \geq \alpha > 1$ ,  $S^*f$  is weak  $L_1$  and strong  $L_p$  for  $1 < p \leq 2$ .*

**Remark 2.12.** As in Remark 2.5(a), this result implies the usual maximal inequalities in the ergodic theorem. Furthermore, with a similar proof, the same result as the one in Theorem 2.11 holds for  $S_M^*f$  for arbitrary  $(n_k)$  and lacunary  $M$  because of Theorem 1.22.

The problem with extending Theorem 2.11 to cover all  $(n_k)$  is that we do not know when the strong  $L_2$  inequality holds. By White's result, if  $n_{k+1} \geq n_k^p$  for some  $p > 1$ , then this is the case. But even in this case, it is not clear when  $S^*f$  is strong  $L_p$ ,  $2 < p < \infty$ . But it is important to remark in this regard that Bourgain [6] has shown

that functions related to the square function are always strong  $L_2$ . For example, using his result and the technique in Theorem 2.10, one can see that for any  $(n_k)$ ,

$$S_4^* f = \left( \sum_{k=1}^{\infty} \sup_{n_k \leq n \leq n_{k+1}} |A_n f - A_{n_k} f|^4 \right)^{1/4}$$

is weak  $(1, 1)$  and strong  $(p, p)$  at least for  $1 < p \leq 2$ .

It is worthwhile to point out that there is always these easier facts about square functions for block maxima. Here we use the usual maximal function  $f^*(x) = \sup_{n \geq 1} |A_n f(x)|$ .

**THEOREM 2.13.** *Let  $(n_k)$  denote an increasing sequence of integers. If  $n_k = p(k)$  for some polynomial  $p$  of degree  $s > 0$ , then there is a constant  $C$  such that  $\|S^* f\|_2 \leq C \|f\|_2$  for all  $f \in L_2(X)$  and consequently  $S^*$  is weak type  $(1, 1)$ . Furthermore, there is a constant  $C_p$ , for  $1 < p \leq \infty$ , such that  $\|S^* f\|_p \leq C_p \|f\|_p$  for all  $f \in L_p(X)$ .*

*Proof.* We have

$$\begin{aligned} S^* f(x) &= \left( \sum_{k=1}^{\infty} \sup_{n_k \leq n \leq n_{k+1}} |A_n f(x) - A_{n_k} f(x)|^2 \right)^{1/2} \\ &\leq \left( \sum_{k=1}^{\infty} \sup_{n_k \leq n \leq n_{k+1}} \left| \left( \frac{1}{n} - \frac{1}{n_k} \right) n_k A_{n_k} f(x) + \frac{1}{n} \sum_{r=n_k}^n f(\tau^r x) \right|^2 \right)^{1/2} \\ &\leq \left( 2 \sum_{k=1}^{\infty} \sup_{n_k \leq n \leq n_{k+1}} \left| \frac{n - n_k}{n} A_{n_k} f(x) \right|^2 + 2 \sum_{k=1}^{\infty} \sup_{n_k \leq n \leq n_{k+1}} \left| \frac{1}{n} \sum_{r=n_k}^n f(\tau^r x) \right|^2 \right)^{1/2} \\ &\leq \left( 2 f^*(x)^2 \sum_{k=1}^{\infty} \sup_{n_k \leq n \leq n_{k+1}} \left| \frac{n - n_k}{n} \right|^2 + 2 \sum_{k=1}^{\infty} \sup_{n_k \leq n \leq n_{k+1}} \left| \frac{1}{n} \sum_{r=n_k}^n f(\tau^r x) \right|^2 \right)^{1/2} \\ &\leq \left( 2 f^*(x)^2 \sum_{k=1}^{\infty} \left( \frac{n_{k+1} - n_k}{n_k} \right)^2 + 2 \sum_{k=1}^{\infty} \left( \frac{1}{n_k} \sum_{r=n_k}^{n_{k+1}} |f(\tau^r x)| \right)^2 \right)^{1/2} \\ &\leq \left( 2 f^*(x)^2 \sum_{k=1}^{\infty} \left( \frac{C}{k} \right)^2 + \sum_{k=1}^{\infty} \left( \frac{C}{k} \frac{1}{k^{s-1}} \sum_{r=n_k}^{n_{k+1}} |f(\tau^r x)| \right)^2 \right)^{1/2} \\ &\leq c f^*(x) + \left( \sum_{k=1}^{\infty} \frac{C}{k^2} \left( \frac{1}{k^{s-1}} \sum_{r=n_k}^{n_{k+1}} |f(\tau^r x)| \right)^2 \right)^{1/2} \end{aligned}$$

Thus the result will follow if we can show that the operator

$$\tilde{S} f(x) = \left( \sum_{k=1}^{\infty} \frac{1}{k^2} \left( \frac{1}{k^{s-1}} \sum_{r=n_k}^{n_{k+1}} |f(\tau^r x)| \right)^2 \right)^{1/2}$$

is strong type  $(2, 2)$ . This follows easily by just integrating and interchanging the order of integration and summation. To see  $S^*$  is strong  $(p, p)$ , use the above computation and observe that  $f^*$  and  $\tilde{S}$  are bounded operators on  $L_{\infty}$ . Then interpolate using the weak type  $(1, 1)$  that follows by the above and Theorem 2.10.  $\square$

**Remark 2.14.** The obvious conjecture from all the above, is that for any  $(n_k)$  increasing,  $S^*$  is weak  $(1, 1)$  and strong  $(p, p)$ ,  $1 < p < \infty$ .

### 3. Square functions and random translations

There is an interesting aspect of square functions that is especially useful in ergodic theory: for strong  $L_2$  estimates, the terms may be translated randomly. Indeed, fix  $(n_k)$ ,  $n_{k+1} \geq n_k$ , and  $(m_k)$ . Then

$$\left\| \left( \sum_{k=1}^{\infty} |(A_{n_{k+1}} - A_{n_k})f \circ \tau^{m_k}|^2 \right)^{1/2} \right\|_2 = \left\| \left( \sum_{k=1}^{\infty} |(A_{n_{k+1}} - A_{n_k})f|^2 \right)^{1/2} \right\|_2.$$

This fact and the transfer methods from  $(X, \beta, m, \tau)$  to  $(\mathbb{Z}, +1)$  give this corollary to Theorem 1.1.

**THEOREM 3.1.** *For any  $(n_k)$ ,  $n_{k+1} \geq n_k$  for all  $k \geq 1$ , and for any  $(m_k)$ ,*

$$\left\| \sup_k |(A_{n_{k+1}} - A_{n_k})f \circ \tau^{m_k}| \right\|_2 \leq 25 \|f\|_2.$$

There is also another immediate corollary of bounds for the square function.

**COROLLARY 3.2.** *Let  $(n_k)$  be arbitrary,  $n_k \leq n_{k+1}$  for all  $k \geq 1$ . Then*

$$\sum_{k=1}^{\infty} m\{|(A_{n_{k+1}} - A_{n_k})f| > \lambda\} \leq \frac{625}{\lambda^2} \|f\|_2^2$$

*for all dynamical systems  $(X, \beta, m, \tau)$ .*

*Proof.* Clearly,

$$\begin{aligned} \sum_{k=1}^{\infty} m\{|(A_{n_{k+1}} - A_{n_k})f| > \lambda\} &\leq \frac{1}{\lambda^2} \sum_{k=1}^{\infty} \|(A_{n_{k+1}} - A_{n_k})f\|_2^2 \\ &= \frac{1}{\lambda^2} \left\| \left( \sum_{k=1}^{\infty} |(A_{n_{k+1}} - A_{n_k})f|^2 \right)^{1/2} \right\|_2^2 \\ &\leq \frac{625}{\lambda^2} \|f\|_2^2. \end{aligned}$$

□

The important point to be made here is that the condition of Theorem 3.1 is equivalent to Theorem 1.1, and the constant does not need to be independent of  $(m_k)$ . This follows from the following theorem.

**THEOREM 3.3.** *Let  $(d_k)$  be a sequence of finite measures on  $\mathbb{Z}$ . Then the following are equivalent:*

(1) *there is a constant  $C$  with*

$$\left\| \left( \sum_{k=1}^{\infty} |d_k * \varphi|^2 \right)^{1/2} \right\|_{\ell_2} \leq C \|\varphi\|_{\ell_2};$$

(2) *there is a constant  $C$  such that for all  $(m_k)$ ,*

$$\left\| \sup_k |d_k * \delta_{m_k} * \varphi| \right\|_{\ell_2} \leq C \|\varphi\|_{\ell_2};$$

(3) for each sequence  $(m_k)$ , there is a constant  $C$  such that

$$\left\| \sup_k |d_k * \delta_{m_k} * \varphi| \right\|_{\ell_2} \leq C \|\varphi\|_{\ell_2}.$$

*Proof.* Clearly (1) implies (2) and (2) implies (3). Also, each of (1), (2), and (3) holds for  $(d_k)$  if and only if it holds for any sequence of finite measures  $d'_k$  with  $\sum_{k=1}^{\infty} \|d_k - d'_k\|_{\ell_1} < \infty$ . So without loss of generality, we can assume the  $(d_k)$  have finite support. Assume (2) and that  $\varphi$  has finite support. Choose  $(m_k)$  so that  $(d_k * \delta_{m_k} * \varphi)$  are disjointly supported; then,

$$\left( \sum_{k=1}^{\infty} |d_k * \delta_{m_k} * \varphi|^2 \right)^{1/2} = \sup_k |d_k * \delta_{m_k} * \varphi|.$$

Hence,

$$\begin{aligned} \left\| \left( \sum_{k=1}^{\infty} |d_k * \varphi|^2 \right)^{1/2} \right\|_{\ell_2}^2 &= \sum_{k=1}^{\infty} \|d_k * \varphi\|_{\ell_2}^2 \\ &= \sum_{k=1}^{\infty} \|d_k * \delta_{m_k} * \varphi\|_{\ell_2}^2 \\ &= \left\| \left( \sum_{k=1}^{\infty} |d_k * \delta_{m_k} * \varphi|^2 \right)^{1/2} \right\|_{\ell_2}^2 \\ &= \left\| \sup_k |d_k * \delta_{m_k} * \varphi| \right\|_{\ell_2}^2 \\ &\leq C^2 \|\varphi\|_{\ell_2}^2. \end{aligned}$$

Hence, (1) holds for all  $\varphi$  with finite support. A routine approximation argument proves (1) holds for all  $\varphi$ .

Now assume (3) and fix  $0 \neq \varphi \in \ell_2$ . Let  $\varepsilon_k > 0$  with  $\sum_{k=1}^{\infty} \varepsilon_k^2 \leq \|\varphi\|_{\ell_2}^2$  and choose finite sets  $E_k \subset \mathbb{Z}$  with  $\|(d_k * \varphi)1_{E_k} - d_k * \varphi\|_2 \leq \varepsilon_k$ . Then choose  $(\delta_{m_k})$  such that  $((d_k * \varphi)1_{E_k}) * \delta_{m_k}$  are pairwise disjointly supported; then,

$$\sum_{k=1}^{\infty} \|(d_k * \varphi)1_{E_k} * \delta_{m_k}\|_{\ell_2}^2 = \left\| \sup_k |(d_k * \varphi)1_{E_k} * \delta_{m_k}| \right\|_{\ell_2}^2.$$

We have

$$\begin{aligned} \left\| \left( \sum_{k=1}^{\infty} |d_k * \varphi|^2 \right)^{1/2} \right\|_{\ell_2}^2 &= \sum_{k=1}^{\infty} \|d_k * \varphi\|_{\ell_2}^2 \\ &\leq 2 \sum_{k=1}^{\infty} (\|(d_k * \varphi)1_{E_k}\|_{\ell_2}^2 + \varepsilon_k^2) \\ &= 2 \sum_{k=1}^{\infty} (\|(d_k * \varphi)1_{E_k} * \delta_{m_k}\|_{\ell_2}^2 + \varepsilon_k^2) \\ &\leq 2\|\varphi\|_{\ell_2}^2 + 2 \sum_{k=1}^{\infty} \|(d_k * \varphi)1_{E_k} * \delta_{m_k}\|_{\ell_2}^2 \end{aligned}$$

$$\begin{aligned}
 &= 2\|\varphi\|_{\ell_2}^2 + 2\left\|\sup_k |(d_k * \varphi)1_{E_k} * \delta_{m_k}|\right\|_{\ell_2}^2 \\
 &\leq 2\|\varphi\|_{\ell_2}^2 + 4\left\|\sup_k |d_k * \varphi * \delta_{m_k}|\right\|_{\ell_2}^2 \\
 &\quad + 4\left\|\sup_k |(d_k * \varphi - (d_k * \varphi)1_{E_k}) * \delta_{m_k}|\right\|_{\ell_2}^2 \\
 &\leq 2\|\varphi\|_{\ell_2}^2 + 4C\|\varphi\|_{\ell_2}^2 + 4\sum_{k=1}^{\infty} \|d_k * \varphi - (d_k * \varphi)1_{E_k}\|_{\ell_2}^2 \\
 &\leq 2\|\varphi\|_{\ell_2}^2 + 4C\|\varphi\|_{\ell_2}^2 + 4\sum_{k=1}^{\infty} \varepsilon_k^2 \\
 &\leq 2\|\varphi\|_{\ell_2}^2 + 4C\|\varphi\|_{\ell_2}^2 + 4\|\varphi\|_{\ell_2}^2 \\
 &= C\|\varphi\|_{\ell_2}^2
 \end{aligned}$$

for some constant  $C$ , depending on  $(m_k)$ . □

**COROLLARY 3.4.** *For a sequence of finite measures  $(d_k)$  on  $\mathbb{Z}$ , consider:*

- (1) *for some constant  $C$ ,  $\|(\sum_{k=1}^{\infty} |d_k f|^2)^{1/2}\|_{L_2} \leq C\|f\|_{L_2}$ ;*
- (2) *for some constant  $C$ ,  $\|\sup_{k \geq 1} |d_k f \circ \tau^{m_k}|\|_{L_2} \leq C\|f\|_{L_2}$  for all  $(m_k)$ ;*
- (3) *for each  $(m_k)$ ,  $\|\sup_{k \geq 1} |d_k f \circ \tau^{m_k}|\|_{L_2} \leq C\|f\|_{L_2}$  for some constant  $C$ .*

*If (1), (2), or (3) holds for some aperiodic non-atomic finite dynamical system, then they all hold for all dynamical systems.*

*Proof.* If (1), (2), or (3) hold for some aperiodic non-atomic system, then by using the Rokhlin lemma, it is easy to see (1), (2), or (3) respectively of Theorem 3.3 holds. So by the Calderón transfer principle, the proof is complete. □

There are many alternative versions of the idea above. Here is a particular one that is representative of this. We state this principle for  $(\mathbb{Z}, +1)$ , but it holds in any dynamical system; indeed, either property holds in  $(\mathbb{Z}, +1)$  if and only if it holds for some (or all) aperiodic dynamical systems.

**THEOREM 3.5.** *For a sequence of finite measures  $(d_k)$  on  $\mathbb{Z}$ , the following are equivalent for  $1 \leq p \leq \infty$ :*

- (1) *for some constant  $C$ ,*

$$\left\|\left(\sum_{k=1}^{\infty} |d_k * \varphi \circ \delta_{m_k}|^2\right)^{1/2}\right\|_{\ell_p} \leq C\|\varphi\|_{\ell_p}$$

*for all  $(m_k)$ ;*

- (2) *for each  $(m_k)$ , there is some constant  $C$*

$$\left\|\left(\sum_{k=1}^{\infty} |d_k * \varphi * \delta_{m_k}|^2\right)^{1/2}\right\|_{\ell_p} \leq C\|\varphi\|_{\ell_p}.$$

*Proof.* Clearly (1) implies (2).

Assume (2) and that (1) fails to hold, then we can inductively choose  $(\varphi_j)$ ,  $\|\varphi_j\|_{\ell_p}$  as small as we like, with

$$\left\| \left( \sum_{k=1}^{N_j} |d_k * \varphi_j * \delta_{m'_k}|^2 \right)^{1/2} \right\|_{\ell_p}$$

as large as we like for suitable  $(m'_j)$ . To be explicit, fix  $\varepsilon_k > 0$  with  $\sum_{k=1}^{\infty} \varepsilon_k \leq 1$  and let  $\|\varphi_1\|_{\ell_p} \leq \varepsilon_1$ , but for suitable  $N_1, m'_1, \dots, m'_{N_1}$ ,

$$\left\| \left( \sum_{k=1}^{N_1} |d_k * \varphi_1 * \delta_{m'_k}|^2 \right)^{1/2} \right\|_{\ell_p} \geq K_1.$$

Now choose  $\varphi_2$  with  $\|\varphi_2\|_{\ell_p} \leq \varepsilon_2$  and for some  $N_2, m_1^2, \dots, m_{N_2}^2$ ,

$$\left\| \left( \sum_{k=N_1+1}^{N_2} |d_k * \varphi_2 * \delta_{m_k^2}|^2 \right)^{1/2} \right\|_{\ell_p} \geq K_2.$$

This is possible because

$$\left\| \left( \sum_{k=1}^{N_1} |d_k * \varphi_2 * \delta_{m_k^2}|^2 \right)^{1/2} \right\|_{\ell_p} \leq \sum_{k=1}^{N_1} \|d_k\|_1 \|\varphi_2\|_{\ell_p},$$

which can be made small by decreasing  $\|\varphi_2\|_{\ell_p}$ , independent of the choice of  $(m_k^2)$ . Continue this inductively to generate  $\varphi = \sum_{j=1}^{\infty} \varphi_j \in \ell_p$  and  $(m_k) = (m_1^1, \dots, m_{N_1}^1, m_1^2, \dots, m_{N_2}^2, \dots)$  with

$$\left\| \left( \sum_{k=N_{j-1}+1}^{N_{j+1}} |d_k * \varphi_{j+1} * \delta_{m_k^{j+1}}|^2 \right)^{1/2} \right\|_{\ell_p} \geq K_{j+1}.$$

But now

$$\begin{aligned} & \left\| \left( \sum_{k=1}^{N_{j+1}} |d_k * \varphi * \delta_{m_k}|^2 \right)^{1/2} \right\|_{\ell_p} \\ & \geq \left\| \left( \sum_{k=1}^{N_{j+1}} |d_k * \varphi_{j+1} * \delta_{m_k}|^2 \right)^{1/2} \right\|_{\ell_p} - \sum_{s=1}^j \left\| \left( \sum_{k=1}^{N_{j+1}} |d_k * \varphi_s * \delta_{m_k}|^2 \right)^{1/2} \right\|_{\ell_p} \\ & \quad - \sum_{s=j+2}^{\infty} \left\| \left( \sum_{k=1}^{N_{j+1}} |d_k * \varphi_s * \delta_{m_k}|^2 \right)^{1/2} \right\|_{\ell_p} \\ & \geq \left\| \left( \sum_{k=N_j+1}^{N_{j+1}} |d_k * \varphi_{j+1} * \delta_{m_k}|^2 \right)^{1/2} \right\|_{\ell_p} - \sum_{s=1}^j \left\| \left( \sum_{k=1}^{N_{j+1}} |d_k * \varphi_s * \delta_{m_k}|^2 \right)^{1/2} \right\|_{\ell_p} \\ & \quad - \sum_{s=j+2}^{\infty} \left\| \left( \sum_{k=1}^{N_{j+1}} |d_k * \varphi_s * \delta_{m_k}|^2 \right)^{1/2} \right\|_{\ell_p} \\ & \geq K_{j+1} - C \sum_{s=1}^j \|\varphi_s\|_{\ell_p} - \sum_{s=j+2}^{\infty} \sum_{k=1}^{N_{j+1}} \|d_k\|_{\ell_1} \|\varphi_s\|_{\ell_p}, \end{aligned}$$

where  $C$  is the constant guaranteed by (2), which depends on  $(m_k)$ . This shows

$$\left\| \left( \sum_{k=1}^{N_{j+1}} |d_k * \varphi * \delta_{m_k}|^2 \right)^{1/2} \right\|_{\ell_p} \geq K_{j+1} - C \sum_{s=1}^j \varepsilon_s - \sum_{s=j+2}^{\infty} \varepsilon_s \sum_{k=1}^{N_{j+1}} \|d_k\|_{\ell_1}.$$

Clearly, by an inductive choice of  $(\varepsilon_j)$  going to zero rapidly, and  $(K_j)$  going to infinity rapidly, this underestimate can be made to tend to  $\infty$ . But then  $\varphi \in \ell_p$  and

$$\left\| \left( \sum_{k=1}^{\infty} |d_k * \varphi * \delta_{m_k}|^2 \right)^{1/2} \right\|_{\ell_p} = \infty$$

contradicting (2). So (2) implies (1).  $\square$

Now a particular consequence of the idea of random translations is the following theorem.

**THEOREM 3.6.** *Let  $(n_k)$  be an arbitrary non-decreasing sequence and let  $2 \leq p < \infty$ . Then there is a constant  $C_p$  depending only on  $p$  such that*

$$\sum_{k=1}^{\infty} m\{|(A_{n_{k+1}} - A_{n_k})f| > \lambda\} \leq \frac{C_p^p}{\lambda^p} \|f\|_{L_p}^p,$$

for all dynamical systems  $(X, \beta, m, \tau)$ .

*Proof.* Fixing  $(n_k)$ , we take  $d_k = a_{n_{k+1}} - a_{n_k}$ . Then as in Theorem 3.1, Theorem 1.2 implies that for arbitrary  $(m_k)$ ,  $\sup_{k \geq 1} |d_k * \varphi * \delta_{m_k}|$  satisfies a strong  $\ell_2$  maximal inequality. It also clearly satisfies a strong  $\ell_\infty$  maximal inequality. So it satisfies a strong  $\ell_p$  maximal inequality for  $2 \leq p \leq \infty$ . As in Theorem 3.3, the constant in this strong  $\ell_p$  maximal inequality does not depend on  $(m_k)$ , although the independence of the constant is fairly obvious in this case anyway. So for fixed  $p$ ,  $2 \leq p \leq \infty$ , for some constant  $C_p$ , we have

$$\# \left\{ \sup_{k \geq 1} |d_k * \varphi * \delta_{m_k}| > \lambda \right\} \leq \frac{C_p^p}{\lambda^p} \|\varphi\|_{\ell_p}^p.$$

Assume  $\varphi$  is finitely supported. Then with suitable  $(m_k)$ ,  $d_k * \varphi * \delta_{m_k}$  would be disjointly supported. Hence, in this case

$$\begin{aligned} \sum_{k=1}^{\infty} \# \{|d_k * \varphi * \delta_{m_k}| > \lambda\} &= \# \left\{ \sup_{k \geq 1} |d_k * \varphi * \delta_{m_k}| > \lambda \right\} \\ &\leq \frac{C_p^p}{\lambda^p} \|\varphi\|_{\ell_p}^p. \end{aligned}$$

But the left-hand side of this inequality does not depend on  $(m_k)$ . So

$$\sum_{k=1}^{\infty} \# \{|d_k * \varphi| > \lambda\} \leq \frac{C_p^p}{\lambda^p} \|\varphi\|_p^p.$$

An approximation argument gives the same for all  $\varphi$ . Now we use the Calderón transfer principle to transfer this to any dynamical system.  $\square$

The technique above can be used to some limited extent to extend the estimates for square functions to other powers. Namely, let

$$S_q f = \left( \sum_{k=1}^{\infty} |(A_{n_{k+1}} - A_{n_k}) f|^q \right)^{1/q},$$

for  $f \in L_1$ .

**COROLLARY 3.7.** *For arbitrary increasing  $(n_k)$  and for any  $q$ ,  $2 \leq q < \infty$ ,  $S_q f$  is strong  $L_p$  for all  $p$ ,  $2 \leq p \leq q$ .*

*Proof.* The proof of Theorem 3.6 shows for all  $q$ ,  $2 \leq q < \infty$ ,

$$\left\| \sup_k (d_k * \varphi * \delta_{m_k}) \right\|_{\ell_q} \leq C_q \|\varphi\|_{\ell_q},$$

where  $C_q$  does not depend on  $(m_k)$ . If  $\varphi$  has finite support and  $(m_k)$  is properly chosen, then  $(d_k * \varphi * \delta_{m_k})$  are disjointly supported. Hence, then

$$\sup_k |d_k * \varphi * \delta_{m_k}| = \sum_{k=1}^{\infty} |d_k * \varphi * \delta_{m_k}| = \left( \sum_{k=1}^{\infty} |d_k * \varphi * \delta_{m_k}|^q \right)^{1/q}.$$

So

$$\begin{aligned} \left\| \sup_k |d_k * \varphi * \delta_{m_k}| \right\|_{\ell_q}^q &= \left\| \left( \sum_{k=1}^{\infty} |d_k * \varphi * \delta_{m_k}|^q \right)^{1/q} \right\|_{\ell_q}^q \\ &= \sum_{k=1}^{\infty} \| |d_k * \varphi * \delta_{m_k}|^q \|_{\ell_1} \\ &= \sum_{k=1}^{\infty} \| |d_k * \varphi|^q \|_{\ell_1} = \| S_q \varphi \|_{\ell_q}^q. \end{aligned}$$

So  $\|S_q \varphi\|_{\ell_q}^q \leq C_q^q \|\varphi\|_{\ell_q}^q$  for such  $\varphi$ . By approximation, this follows with the same constant for all  $\varphi$ . Now use the Calderón transfer principle to transfer this to any dynamical system.

But in addition, since  $2 \leq q \leq \infty$ , then  $S_q f \leq S_2 f$ . Hence, the strong  $L_2$  estimate for  $S_2 f$  of Theorem 1.1 and interpolation gives the existence of a strong  $L_p$  inequality for  $S_q f$  for all  $p$ ,  $2 \leq p \leq q$ .  $\square$

The above results for large deviations of differences and strong inequalities for  $q$ -functions cannot be generally extended to  $1 \leq p < 2$ , even for lacunary  $(n_k)$ . The example that we present here that shows this is due to Michael Lacey and Maté Wierdl.

**Example 3.8.** Let  $r_k$  denote the  $k$ th Rademacher function,

$$r_k(x) = \begin{cases} 1 & \text{if } x \in [q/2^k, (q+1)/2^k) \text{ for even } q, 0 \leq q < 2^k \\ -1 & \text{otherwise.} \end{cases}$$

and, for a large integer  $N$ , let the function  $f : [0, 1) \rightarrow \mathbb{R}$  be defined by

$$f = r_N + r_{2N} + r_{3N} + \cdots + r_{N \cdot N}.$$



Let  $m$  be the Lebesgue measure and let  $D_k$  denote the Lebesgue derivative

$$D_k f(x) = 2^k \cdot \int_0^{2^{-k}} f(x+y) dm(y).$$

We are going to show that

$$\sum_{n \leq N} m\{|D_{n \cdot N} f(x) - D_{n \cdot N-1} f(x)| \geq 1/2\} \geq \frac{1}{3} \sqrt{N} \cdot \|f\|_{L_1}.$$

The proof of this inequality depends on the following three lemmas. In all three lemmas, we assume that  $x \in [0, 1 - 1/2^{N-1}]$ . This is so that the averages formed by  $D_k$  for  $k \geq N-1$  will stay in the unit interval.

LEMMA 3.8(a) *Let  $\ell > k \geq N$ . Then*

$$D_k r_\ell(x) = 0.$$

*Proof.* This is because in computing  $D_k r_\ell(x)$ , we average on an interval of length  $2^{-k}$ , and  $2^{-k}$  is an integer multiple of  $2 \cdot 2^{-\ell}$ , the period of  $r_\ell$ .  $\square$

LEMMA 3.8(b) *Suppose  $k > \ell \geq N$ , and that  $x$  is not in an interval of the form*

$$\left( \frac{q}{2^\ell} - \frac{1}{2^k}, \frac{q}{2^\ell} \right).$$

*Then*

$$D_k r_\ell(x) = r_\ell(x).$$

*Proof.* This fact holds because in computing  $D_k r_\ell(x)$ , we average on an interval (of length  $2^{-k}$ ) on which  $r_\ell$  is constant (+1 or -1).  $\square$

LEMMA 3.8(c) *Suppose  $k \geq N$ , and that  $x$  is not in the middle half of an interval of the form  $[q/2^k, (q+1)2^k]$ ; that is,  $x$  is not in an interval of the form*

$$\left[ (q+1/4)\frac{1}{2^k}, (q+3/4)\frac{1}{2^k} \right].$$

*Then*

$$|D_k r_k(x)| \geq \frac{1}{2}.$$

*Proof.* Just note that the function  $g(y) = D_k r_k(y)$  is linear on  $[q/2^k, (q+1)2^k]$ , and either  $g(q/2^k) = 1$  and  $g((q+1)2^k) = -1$ , or  $g(q/2^k) = -1$  and  $g((q+1)2^k) = 1$ .  $\square$

We now remove certain ‘bad’ sets from the interval  $[0, 1]$ . (In fact, we remove sets on which the oscillations are potentially small—so from another viewpoint, this is the good set.) Our first bad set does not depend on  $n$ ; it is

$$E = \left( 1 - \frac{1}{2^{N-1}}, 1 \right) \cup \left( \bigcup_{n=1}^N \bigcup_{q=1}^{2^{n-N}} \left( \frac{q}{2^{n \cdot N}} - \frac{1}{2^{(n+1) \cdot N-1}}, \frac{q}{2^{n \cdot N}} \right) \right).$$

We easily get the estimate

$$m(E) \leq 4 \cdot \frac{N}{2^N},$$

so the measure of  $E$  is as small as we want.

For each  $n$ , we further remove a ‘bad’ set; let

$$I_n = \bigcup_{q=1}^{2^{n \cdot N}} \left[ \frac{q + 1/4}{2^{n \cdot N}}, \frac{q + 3/4}{2^{n \cdot N}} \right].$$

Then  $m(I_n) = 1/2$  and hence the measure of the ‘good’ set

$$G = [0, 1] \setminus (E \cup I_n)$$

is as close to  $1/2$  as we want—and certainly greater than  $1/3$ . It remains to show that if  $x \in G$  then

$$|D_{n \cdot N} f(x) - D_{n \cdot N - 1} f(x)| \geq 1/2.$$

Indeed, by Lemmas 3.8(a) and 3.8(b), we have for  $x \in G$ ,

$$D_{n \cdot N} f(x) = \sum_{u < n} r_{u \cdot N}(x) + D_{n \cdot N} r_{n \cdot N}(x),$$

and

$$D_{n \cdot N - 1} f(x) = \sum_{u < n} r_{u \cdot N}(x).$$

Hence, by Lemma 3.8(c), we have for  $x \in G$ ,

$$|D_{n \cdot N} f(x) - D_{n \cdot N - 1} f(x)| = |D_{n \cdot N} r_{n \cdot N}(x)| \geq 1/2.$$

But this inequality says that

$$\sum_{n \leq N} m\{|D_{n \cdot N} f(x) - D_{n \cdot N - 1} f(x)| \geq 1/2\} \geq \frac{1}{3}N.$$

Since  $\|f\|_{L_2} = \sqrt{N}$ , we have by Hölder’s inequality

$$\sum_{n \leq N} m\{|D_{n \cdot N} f(x) - D_{n \cdot N - 1} f(x)| \geq 1/2\} \geq \frac{1}{3} \sqrt{N} \cdot \|f\|_{L^1}.$$

There is also the  $L_p$  result: for  $1 \leq p < 2$ , we have by Hölder’s inequality

$$\sum_{n \leq N} m\{|D_{n \cdot N} f(x) - D_{n \cdot N - 1} f(x)| \geq 1/2\} \geq \frac{1}{3} \cdot N^{1-p/2} \|f\|_{L_p}^p.$$

Although this computation was done in  $[0, 1]$ , it could equally well be done in  $\mathbb{Z}$  and this would deny the analogous inequalities in any ergodic dynamical system. That is, using Proposition 3.8 in Rosenblatt and Wierdl [16], this estimate shows that for any ergodic dynamical system, and any  $p$ ,  $1 \leq p < 2$ , there exists a function  $f \in L_p$  such that

$$\sum_{k=1}^{\infty} m\{|(A_{2^{k+1}} - A_{2^k})f| > 1\} = \infty.$$

In addition, using random translations and Sawyer’s principle, as described in Lemma 2.8 in [16], one can show for any ergodic dynamical system, and any  $p$ ,  $1 \leq p < 2$ , there exists a sequence  $(m_k)$  and a function  $f \in L_p$  such that

$$\sup_k |(A_{2^{k+1}} - A_{2^k})f \circ \tau^{m_k}| = \infty \text{ a.e.}$$

Moreover, in contrast to the positive results which were proved in Theorem 3.7, the associated  $q$ -functions cannot be bounded here. That is, for any ergodic dynamical system, and any  $q$ ,  $1 \leq q < 2$ , and  $p$ ,  $1 \leq p < \infty$ , there exists a function  $f \in L_p$  such that  $S_q f = \infty$  a.e.

Corollary 3.7 and Example 3.8 suggest several very interesting questions concerning large deviations of differences. The most obvious one is what conditions are needed on  $(n_k)$  for there to be a large deviation result in some  $L_p$  as in Corollary 3.7. But more specifically, the question is when does the randomly translated maximal function

$$D^* f = \sup_k |(A_{n_{k+1}} - A_{n_k})f \circ \tau^{m_k}|$$

satisfy a weak  $L_1$ -estimate. If it does, then it is strong  $L_p$ ,  $1 < p < \infty$ , by Marcinkiewicz interpolation. It would follow by Theorem 3.2, that Theorem 1.1 holds, a significantly different approach than the spectral method used previously. Moreover, we can see if  $D^* f$  is weak  $L_1$ , then the constant does not depend on  $(m_k)$  and there is a large deviation inequality:

$$\sum_{k=1}^{\infty} m\{|(A_{n_{k+1}} - A_{n_k})f| > \lambda\} \leq \frac{C}{\lambda} \|f\|_1.$$

Example 3.8 is showing that for lacunary  $(n_k)$  there are no such results. For this reason, the following from [15] is worth pointing out here.

**THEOREM 3.9.** *There is a constant  $C$  such that for all dynamical systems,*

$$\sum_{n=1}^{\infty} m\{|(A_{n+1} - A_n)f| > \lambda\} \leq \frac{C}{\lambda} \|f\|_1.$$

*Proof.* Clearly,

$$|(A_{n+1} - A_n)f| \leq \frac{1}{n+1} A_n |f| + \frac{1}{n+1} |f \circ \tau^{n+1}|.$$

Hence, by a theorem in Rosenblatt and Wierdl [16],

$$\begin{aligned} & \sum_{n=1}^{\infty} m\{|(A_{n+1} - A_n)f| > \lambda\} \\ & \leq \sum_{n=1}^{\infty} m\left\{\frac{1}{n+1} A_n |f| > \frac{\lambda}{2}\right\} + \sum_{n=1}^{\infty} m\left\{\frac{1}{n+1} |f \circ \tau^{n+1}| > \frac{\lambda}{2}\right\} \\ & = \sum_{n=1}^{\infty} m\left\{A_n |f| > \frac{(n+1)\lambda}{2}\right\} + \sum_{n=1}^{\infty} m\left\{|f| > \frac{(n+1)\lambda}{2}\right\} \\ & \leq \frac{C}{\lambda} \|f\|_1. \end{aligned}$$

□

**Remarks 3.10.** The strong  $L_p$  estimate for  $\sup_k |(A_{k+1} - A_k)f \circ \tau^{m_k}|$  when  $1 < p < \infty$  that is a consequence of Theorem 3.5 is trivial since  $\|A_{k+1} - A_k\|_p \leq C/k$  and so trivially

$$\left\| \sup_k |(A_{k+1} - A_k)f \circ \tau^{m_k}| \right\|_p^p \leq \sum_{k=1}^{\infty} \|(A_{k+1} - A_k)f \circ \tau^{m_k}\|_p^p$$

$$\begin{aligned}
&= \sum_{k=1}^{\infty} \|(A_{k+1} - A_k)f\|_p^p \\
&\leq C \sum_{k=1}^{\infty} \frac{1}{k^p} \|f\|_p^p \\
&= C_p \|f\|_p^p.
\end{aligned}$$

The fact that  $C_p$  does not need to tend to  $\infty$  as  $p \rightarrow 1$  is the only point given by Theorem 3.5 for  $L_p$ ,  $p > 1$ .

The same argument as in Theorem 3.5 actually gives the following.

**THEOREM 3.11.** *Let  $n_k = k^L$  for some fixed  $L \in \mathbb{Z}^+$ . Then for some constant  $C_1$ ,*

$$\sum_{k=1}^{\infty} m \left\{ \sup_{n_k \leq n \leq n_{k+1}} |(A_n - A_{n_k})f| > \lambda \right\} \leq \frac{C}{\lambda} \|f\|_1,$$

for all  $f \in L_1$ .

*Proof.* This follows immediately from a result in Rosenblatt and Wierdl [16] and the inequality that for  $n_k \leq n \leq n_{k+1}$ ,

$$\begin{aligned}
|(A_n - A_{n_k})f| &\leq \frac{n - n_k}{n} A_{n_k} |f| + \frac{1}{n} \sum_{\ell=n_k+1}^n |f| \circ \tau^\ell \\
&\leq \frac{n_{k+1} - n_k}{n_{k+1}} A_{n_k} |f| + \frac{1}{n_k} \sum_{\ell=n_k+1}^{n_{k+1}} |f| \circ \tau^\ell \\
&\leq C_L \frac{1}{k} A_{n_k} |f| + \frac{n_{k+1} - n_k}{n_k} A_{n_{k+1} - n_k} |f| \circ \tau^{n_k} \\
&\leq C_L \frac{1}{k} A_{n_k} |f| + C_L \frac{1}{k} A_{n_{k+1} - n_k} |f| \circ \tau^{n_k}.
\end{aligned}$$

□

Let us make one last observation about weak  $L_1$  inequalities for  $D^*$  that was commented on before in a special case.

**THEOREM 3.12.** *For a sequence  $(d_k)$  of finite measures  $\mathbb{Z}$ , the following are equivalent:*

(1) *there is a constant  $C$  such that*

$$\# \left\{ \sup_k |d_k * \varphi * \delta_{m_k}| > \lambda \right\} \leq \frac{C}{\lambda} \|\varphi\|_{\ell_1}$$

for all  $(m_k)$ ;

(2) *for each  $(m_k)$ , there is a constant  $C$  with*

$$\# \left\{ \sup_k |d_k * \varphi * \delta_{m_k}| > \lambda \right\} \leq \frac{C}{\lambda} \|\varphi\|_{\ell_1},$$

(3)

$$\sum_{k=1}^{\infty} \# \{ |d_k * \varphi| > \lambda \} \leq \frac{C}{\lambda} \|\varphi\|_{\ell_1}.$$

*Proof.* Clearly (1) implies (2). Also, (1) implies (3) by separating the supports of  $d_k * \varphi$  with suitable  $m_k$ , in the case of  $\text{supp } (\varphi)$  being finite. Furthermore, (3) clearly implies (1) because

$$\begin{aligned} \# \left\{ \sup_k |d_k * \varphi * \delta_{m_k}| > \lambda \right\} &\leq \sum_{k=1}^{\infty} \# \{|d_k * \varphi * \delta_{m_k}| > \lambda\} \\ &= \sum_{k=1}^{\infty} \# \{|d_k * \varphi| > \lambda\}. \end{aligned}$$

It remains to prove that (2) implies (1). This can be done in a manner similar to the proof in Theorem 3.5. So assume (2) and that (1) fails. Then for each constant  $C$ , and  $\varepsilon > 0$ , there is  $\varphi$ ,  $\|\varphi\|_1 \leq \varepsilon$ , and some  $(m_k)$  such that for some  $\lambda$ ,

$$\sup_k \lambda \# \left\{ \sup_k |d_k * \varphi * \delta_{m_k}| > \lambda \right\} > C.$$

Hence, there is some  $m_1^1, \dots, m_{N_1}^1$  and  $\lambda_1$  with

$$\lambda_1 \# \left\{ \sup_{1 \leq k \leq N_1} |d_k * \varphi * \delta_{m_k^1}| > \lambda_1 \right\} > C.$$

We inductively choose  $C_\ell$ ,  $\varepsilon_\ell$ ,  $\varphi_\ell$  and  $m_1^\ell, \dots, m_{N_\ell}^\ell$  so that  $\|\varphi_\ell\|_{\ell_1} < \varepsilon_\ell$  and for some  $\lambda_\ell > 0$ ,

$$\lambda_\ell \# \left\{ \sup_{N_{\ell-1}+1 \leq k \leq N_\ell} |d_k * \varphi_\ell * \delta_{m_k^\ell}| > \lambda_\ell \right\} > C_\ell.$$

Let  $\varphi = \sum_{\ell=1}^{\infty} \varphi_\ell$ . Now for each  $\ell$ ,  $\ell \geq 2$ ,

$$\begin{aligned} &\lambda_\ell \# \left\{ \sup_{N_{\ell-1}+1 \leq k \leq N_\ell} |d_k * \varphi * \delta_{m_k^\ell}| > \lambda_\ell \right\} \\ &\geq \lambda_\ell \# \left\{ \sup_{N_{\ell-1}+1 \leq k \leq N_\ell} |d_k * \varphi_\ell * \delta_{m_k^\ell}| > 3\lambda_\ell \right\} \\ &\quad - \lambda_\ell \# \left\{ \sup_{N_{\ell-1}+1 \leq k \leq N_\ell} \left| d_k * \sum_{s=1}^{\ell-1} \varphi_s * \delta_{m_k^\ell} \right| > \lambda_\ell \right\} \\ &\quad - \lambda_\ell \# \left\{ \sup_{N_{\ell-1}+1 \leq k \leq N_\ell} \left| d_k * \sum_{s=\ell+1}^{\infty} \varphi_s * \delta_{m_k^\ell} \right| > \lambda_\ell \right\} \\ &\geq \lambda_\ell \# \left\{ \sup_{N_{\ell-1}+1 \leq k \leq N_\ell} |d_k * \varphi_\ell * \delta_{m_k^\ell}| > 3\lambda_\ell \right\} \\ &\quad - \sum_{s=1}^{\ell-1} \lambda_\ell \left\{ \sup_{N_{\ell-1}+1 \leq k \leq N_\ell} |d_k * \varphi_s * \delta_{m_k^\ell}| > \frac{\lambda_\ell}{\ell-1} \right\} \\ &\quad - \lambda_\ell \# \left\{ \sum_{k=N_{\ell-1}+1}^{N_\ell} \left| d_k * \sum_{s=\ell+1}^{\infty} \varphi_s * \delta_{m_k^\ell} \right| > \lambda_\ell \right\} \\ &\geq C_\ell - \sum_{s=1}^{\ell-1} (\ell-1) C \|\varphi_s\|_{\ell_1} - \left\| \sum_{k=N_{\ell-1}+1}^{N_\ell} \left| d_k * \sum_{s=\ell+1}^{\infty} \varphi_s * \delta_{m_k^\ell} \right| \right\|_{\ell_1} \end{aligned}$$

where  $C$  is a constant depending on  $(m_k^s)$  guaranteed to exist by (2). Since

$$\left\| \sum_{k=N_{\ell-1}+1}^{N_{\ell}} \left| d_k * \sum_{s=\ell+1}^{\infty} \varphi_s * \delta_{m_k^s} \right| \right\|_{\ell_1} \leq \sum_{k=N_{\ell-1}+1}^{N_{\ell}} \|d_k\|_{\ell_1} \cdot \sum_{s=\ell+1}^{\infty} \|\varphi_s\|_{\ell_1},$$

by a suitable choice of  $(\varepsilon_{\ell})$ ,  $C_{\ell}$ , we can guarantee that

$$\lambda_{\ell} \# \left\{ \sup_{N_{\ell-1}+1 \leq k \leq N_{\ell}} |d_k * \varphi * \delta_{m_k^{\ell}}| > \lambda_{\ell} \right\}$$

is unbounded as  $\ell$  tends to  $\infty$ . But then with  $\tilde{m}_k$  being  $(m_k^{\ell} : \ell \geq 1, k = 1, \dots, N_{\ell})$ ,

$$\lambda_{\ell} \# \left\{ \sup_k |d_k * \varphi * \delta_{\tilde{m}_k}| > \lambda_{\ell} \right\}$$

is unbounded as  $\ell$  tends to  $\infty$ , so (2) fails for  $\tilde{m}$ . Hence, (2) implies (1).  $\square$

This last result should clarify the connection between any weak  $L_1$  estimate for a randomly translated maximal function and the large deviations of the operators in question. In particular, when both hold,

$$\sum_{k=1}^{\infty} m\{|d_k * f| > \lambda\} < \infty,$$

for all  $\lambda > 0$ ,  $f \in L_1$ . In this context, this itself is usually enough to give the homogeneous inequalities of Theorem 3.12. Indeed, from Theorem 3.12, we see that the issue of proving a large deviation result like

$$\sum_{k=1}^{\infty} m\{|(A_{n_{k+1}} - A_{n_k})f| > \lambda\} \leq \frac{1}{\lambda} \|f\|_1$$

can be equivalently formulated as showing that for any fixed  $(n_k)$ , and  $f \in L_1$ , if  $(m_k)$  as arbitrary (perhaps even rapidly enough increasing) then

$$(A_{n_{k+1}} - A_{n_k})f \circ \tau^{m_k}(x) \rightarrow 0 \text{ a.e. } x.$$

Moreover, by the usual transfer methods, such a result would hold in some aperiodic dynamical system if and only if it held in them all. But the non-homogeneous inequality above would give this convergence result in any dynamical system in which it held.

*Note.* In joint work of R. Jones, R. Kaufman, J. Rosenblatt and M. Wierdl done after this article was in press, a number of the questions unresolved here have been answered; this work will appear soon elsewhere.

*Acknowledgements.* We would like to thank Michael Lacey who suggested the use of abstract Calderón–Zygmund operators in §2, Mate Wierdl who suggested the method of estimating  $\Sigma_2$  used in §1 in place of an earlier technique, and Robert Bradley who pointed out Gaposhkin's results for square functions. We also greatly appreciate the example given to us by Michael Lacey and Maté Wierdl which was included in §3.

## REFERENCES

- [1] 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, 1989)*. Academic Press, New York, 1991, pp. 17–40.
- [2] A. Bellow, R. Jones and J. Rosenblatt. Almost everywhere convergence of weighted averages. *Math. Annalen* **293** (1992), 399–426.
- [3] A. Bellow, R. Jones and J. Rosenblatt. Almost everywhere convergence of convolution powers. *Ergod. Th. & Dynam. Sys.* **14** (1994), 415–432.
- [4] A. Benedek, A. P. Calderón and R. Panzone. Convolution operators on Banach space valued functions. *Proc. National Academy of Sciences USA* **48** (1962), 356–365.
- [5] J. Bourgain. On the maximal ergodic theorem for certain subsets of the integers. *Israel J. Math.* **61** (1988), 39–72.
- [6] J. Bourgain. Pointwise ergodic theorems for arithmetic sets. *Publications Mathématiques, Institut des Hautes Etudes Scientifiques* **69** (1989), 5–45.
- [7] R. E. Bradley. Pseudospectral operators and the pointwise ergodic theorem. *Proc. Amer. Math. Soc.* **112** (1991), 863–870.
- [8] D. L. Burkholder. Martingale transforms. *Ann. Math. Statist.* **37** (1966), 1494–1504.
- [9] D. L. Burkholder, B. J. Davis and R. F. Gundy. Integral inequalities for convex functions of operators on martingales. *Proc. 6th Berkeley Symp. on Mathematical Statistics and Probability 2*. University of California Press, Berkeley, CA, 1972, pp. 223–240.
- [10] A. P. Calderón. Ergodic theory and translation invariant operators. *Proc. National Academy of Sciences USA* **59** (1968) 349–353.
- [11] J. Duoandikoetxea and J. Rubio de Francia. Maximal and singular integral operators via Fourier transform estimates. *Inventiones mathematicae* **84** (1986), 541–561.
- [12] 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.
- [13] V. F. Gaposhkin. Individual ergodic theorem for normal operators on  $L_2$ . *Functional Analysis and Applications* **15** (1981), 14–18.
- [14] V. F. Gaposhkin and J. Rosenblatt. Almost everywhere convergence of sequences of powers. *Proc. 5th Conf. on Probability and Statistics, (Vilnius, 1989)*, Vol I. Mokslas, Utrecht, 1990, pp. 391–400.
- [15] R. Jones. Inequalities for the ergodic maximal function. *Studia Math* **40** (1977), 111–129.
- [16] J. Rosenblatt and M. Wierdl. A new maximal inequality and its applications. *Ergod. Th. & Dynam. Sys.* **12** (1992), 509–558.
- [17] J. Rosenblatt and M. Wierdl. Pointwise ergodic theorems via harmonic analysis. *Proc. Conf. on Ergodic Theory (Alexandria, Egypt, 1993)*. Cambridge University Press, 1994, pp. 3–151.
- [18] J. Rubio de Francia, F. Ruiz and J. Torrea. Calderón–Zygmund theory for operator-valued kernels. *Adv. Math.* **62** (1986), 7–48.
- [19] B. Sz-Nagy and C. Foias. *Analyse harmonique des operateurs de l'espace de Hilbert*. Akad. Kiado, Budapest. Mason, Paris, 1967.
- [20] H. White. The pointwise ergodic theorem and related analytic inequalities. *Master's Thesis*. University of North Carolina, Chapel Hill, NC, 1989.
- [21] M. Wierdl. Almost everywhere convergence and recurrence along subsequences in ergodic theory. *Ph.D. Thesis*. The Ohio State University, 1989.
- [22] M. Wierdl. Pointwise ergodic theorem along the prime numbers. *Israel J. Math.* **64** (1988), 315–336.