Canad. J. Math. Vol. **51** (5), 1999 pp. 996–1019

Counting in Ergodic Theory

Roger L. Jones, Joseph M. Rosenblatt and Máté Wierdl

Abstract. Many aspects of the behavior of averages in ergodic theory are a matter of counting the number of times a particular event occurs. This theme is pursued in this article where we consider large deviations, square functions, jump inequalities and related topics.

1 Introduction

A number of phenomena in ergodic theory can be understood by counting occurrences of various events. In Section 2, counting of levels shows how large deviation theorems immediately imply control of series like square functions. This leads to some unusual rate results for averages somewhat like the ones in the usual ergodic theorem. In Section 3, counting of jumps for martingales leads to a technique to show that the jump inequalities previously obtained for ergodic averages are the best possible ones. See also Ivanov [10] and Kachurovskii [14] for closely related results. In Section 4, counting the leading edge of the ergodic average gives an ergodic theorem due to Assani [2, 3] in L_p for $p > 1$ together with some improvements as a result of the theorems in Section 2. The theme of counting occurrences of various sorts for ergodic averages is seen in all of the sections.

2 Square Functions Via Level Counting

Let (X, β, μ) be a probability space and τ be an invertible measure-preserving transformation of (X, β, μ). Consider a sequence of moving averages $M_n f(x) = \frac{1}{L_n} \sum_{k=v_n+1}^{v_n+L_n} f(\tau^k x)$. It was shown by Rosenblatt and Wierdl [16] that if L_n is non-decreasing, then $M_n f(x)/n$ converges to 0 a.e. $[\mu]$ for any $f \in L_1(X)$. Various ways in which this result is best-possible are discussed in [16], but there still remained the possibility of a result concerning the rate at which $M_n f(x)/n$ converges to 0 in the form of a series condition. For instance, in Jones, Ostrovskii, and Rosenblatt [12], it was asked whether the square function $Sf(x) = (\sum_{n=1}^{\infty} |\frac{M_n f(x)}{n}|^2)^{1/2}$ is finite a.e. for any $f \in L_1(X)$. It turns out that by using the large deviation results in Rosenblatt and Wierdl [16], this can be shown. Let us first state the large deviation result in the form that we will use it. This is Theorem 3.1 in Rosenblatt and Wierdl [16].

Theorem 2.1 *Let (L_n) be a non-decreasing sequence of whole numbers. Let $\lambda > 0$. Then*

$$\sum_{n=1}^{\infty} \mu\{x \in X : M_n f(x) \geq \lambda n|\} \leq \frac{2}{\lambda} \|f\|_1.$$

Received by the editors April 2, 1998; revised February 19, 1999 and April 16,1999.

R. Jones is partially supported by NSF Grant DMS—9531526. J. Rosenblatt is partially supported by NSF Grant DMS—9705228. M. Wierdl is partially supported by NSF Grant DMS—9500577.

AMS subject classification: ****MISSING AMS CLASSIFICATION****.

This result gives the following theorem.

Theorem 2.2 *For any sequence of moving averages $M_n f(x)$ with non-decreasing lengths (L_n), the operator Sf is finite a.e. for all $f \in L_1(X)$, and is weak type $(1,1)$.*

Proof We can assume without loss of generality that $f \in L_1(X)$ is positive a.e. Let us first show that $Sf(x)$ is finite a.e. Write

$$Sf(x)^2 = \sum_{n=1}^{\infty}\Big(\frac{M_n f(x)}{n}\Big(\sum_{m=0}^{\infty} 1_{\{\frac{1}{2^{m+1}} \le \frac{M_n f}{n} < \frac{1}{2^m}\}}(x) + 1_{\{1 \le \frac{M_n f}{n}\}}(x)\Big)\Big)^2.$$

Because of the disjointness of the level sets for $\frac{M_n f}{n}$, we have

$$Sf(x)^2 = \sum_{n=1}^{\infty}\Big(\frac{M_n f(x)}{n}\Big(\sum_{m=0}^{\infty} 1_{\{\frac{1}{2^{m+1}} \le \frac{M_n f}{n} < \frac{1}{2^m}\}}(x)\Big)\Big)^2 + \sum_{n=1}^{\infty}\Big(\frac{M_n f(x)}{n} 1_{\{1 \le \frac{M_n f}{n}\}}(x)\Big)^2.$$

Denote the square roots of these last two terms by $S_1 f(x)$ and $S_2 f(x)$ respectively. First, let us consider $S_2 f(x)$. By Theorem 2.1, we know that $\sum_{n=1}^{\infty} \mu(\{M_n f \ge n\}) < \infty$. That is, the series $\sum_{n=1}^{\infty} 1_{\{1 \le \frac{M_n f}{n}\}}$ is integrable and therefore finite a.e. Hence, a.e. x is in only finitely many of the sets $\{M_n f \ge n\}$. This means that the sum giving $S_2 f(x)$ is a.e. a finite sum and hence $S_2 f(x)$ is finite a.e.

Second, let's consider $S_1 f(x)$. By using the disjointness of the level sets again, we see that

$$S_1 f(x)^2 \le \sum_{n=1}^{\infty}\sum_{m=0}^{\infty}\Big(\frac{1}{2^m}\Big)^2 1_{\{\frac{1}{2^{m+1}} \le \frac{M_n f}{n} < \frac{1}{2^m}\}}(x).$$

Reversing the order of summation and integrating gives

$$\int_X S_1 f(x)^2\, d\mu(x) \le \sum_{m=0}^{\infty} \frac{1}{2^{2m}} \sum_{n=1}^{\infty} \mu\Big(\Big\{\frac{1}{2^{m+1}} \le \frac{M_n f}{n} < \frac{1}{2^m}\Big\}\Big).$$

But $\mu(\{\frac{1}{2^{m+1}} \le \frac{M_n f}{n} < \frac{1}{2^m}\}) \le \mu(\{\frac{1}{2^{m+1}} \le \frac{M_n f}{n}\})$ and so

$$\int_X S_1 f(x)^2\, d\mu(x) \le \sum_{m=0}^{\infty} \frac{1}{2^{2m}} \sum_{n=1}^{\infty} \mu\Big(\Big\{\frac{1}{2^{m+1}} \le \frac{M_n f}{n}\Big\}\Big).$$

By Theorem 2.1, we have the estimate that

$$\sum_{n=1}^{\infty} \mu\Big(\Big\{\frac{1}{2^{m+1}} \le \frac{M_n f}{n}\Big\}\Big) \le 2^{m+1} C \|f\|_1.$$

https://doi.org/10.4153/CJM-1999-044-2 Published online by Cambridge University Press

Hence,

$$\int_X S_1 f(x)^2 \, d\mu(x) \leq C\|f\|_1 \sum_{m=0}^{\infty} \frac{1}{2^{2m}} 2^{m+1}$$
$$= 4C\|f\|_1.$$

Thus, $S_1 f(x)^2$ is integrable and hence finite a.e.

The above argument for the parts $S_1 f(x)$ and $S_2 f(x)$, shows that $Sf(x)$ is finite a.e. for any $f \in L_1(X)$. The weak inequality follows by a result in Assani [2] that the Stein-Sawyer principle extends to certain sublinear operators. Alternatively, let $\big(b(k) : k = 1, 2, 3, \dots\big)$ be a sequence dense in the unit ball of $l_2(\mathbb{Z}^+)$. Consider the linear operators

$$S_{N,k} f(x) = \sum_{n=1}^{N} b_n(k) \frac{M_n f(x)}{n}.$$

The result on $Sf(x)$ above shows that the maximal operator $S^* f(x) = \sup_{N,k} |S_{N,k} f(x)|$ is finite a.e. Hence, when τ is ergodic at least, by the Stein-Sawyer principle (see Garsia [9]) $S^* f$ is weak type $(1, 1)$. This shows that $Sf(x)$ is weak type $(1, 1)$ when τ is ergodic because for any $f \in L_1(X)$, $Sf(x) = S^* f(x)$ a.e. by duality in $l_2(\mathbb{Z}^+)$. The weak type $(1, 1)$ inequality now follows for any τ with a constant independent of τ by the Conze principle. ■

One can take a larger view in the above and state generally a principle that shows how a large deviation result always implies a companion theorem on the behavior of square functions; this approach also avoids the use of the modified version of the Stein-Sawyer principle. We actually will not need very much structure for the mappings T_n at all in this result. We need only to assume that each $T_n \colon L_1(X) \to L_1(X)$ is strongly positive (in the sense that $T_n f \geq 0$ for all $f \in L_1(X)$), and that each T_n is positively homogeneous (in the sense that $T_n(cf) = cT_n f$ for nonnegative c and $f \in L_1(X)$). In this case, we say that T_n is a strongly positive, positively homogeneous operator. For instance, T_n could be the absolute value of some linear operator from $L_1(X) \to L_1(X)$.

Theorem 2.3 *Let $(T_n : n \geq 1)$ be a sequence of strongly positive, positively homogeneous operators and let $Sf(x) = \big(\sum_{n=1}^{\infty} T_n f(x)^2\big)^{\frac{1}{2}}$. Assume that for $f \in L_1(X)$,*

$$\sum_{n=1}^{\infty} \mu(\{|T_n f| \geq 1\}) \leq C\|f\|_1.$$

Then

$$\mu(\{Sf \geq \lambda\}) \leq \frac{10C}{\lambda}\|f\|_1.$$

Hence, for any $f \in L_1(X)$, we have $Sf(x) < \infty$ a.e.

Proof We can assume that $\lambda = 2$ and that f is positive. Let us estimate as

$$\begin{aligned} Sf(x) &\le \Big(\sum_{n=1}^{\infty} \big(T_n f(x)\big)^2 1_{\{T_n f<1\}}(x)\Big)^{\frac{1}{2}} + \Big(\sum_{n=1}^{\infty} \big(T_n f(x)\big)^2 1_{\{T_n f\ge 1\}}(x)\Big)^{\frac{1}{2}} \\ &= S_1(x) + S_2(x). \end{aligned}$$

We need to prove the inequalities

$$\mu\big(\{x : S_1(x) \ge 1\}\big) \le C\|f\|_1, \quad \text{and} \tag{2.1}$$

$$\mu\big(\{x : S_2(x) \ge 1\}\big) \le 4C\|f\|_1. \tag{2.2}$$

The proof of (2.2) is immediate because $S_2(x) = 0$ if $\sum_{n=1}^{\infty} 1_{\{T_n f(x)\ge 1\}} = 0$. So we can estimate using Chebychev's inequality and the assumption of the theorem that

$$\begin{aligned} \mu(\{x : S_2(x) \ge 1\}) &\le \mu(\Big\{\sum_{n=1}^{\infty} 1_{\{T_n f\ge 1\}} \ge 1\Big\}) \\ &\le \sum_{n=1}^{\infty} \mu(\{T_n f \ge 1\}) \\ &\le C\|f\|_1. \end{aligned}$$

As for proving (2.1), let us estimate first by

$$\begin{aligned} S_1(x) &= \Big(\sum_{n=1}^{\infty} \big(T_n f(x)\big)^2 \cdot \sum_{k=0}^{\infty} 1_{\{\frac{1}{2^{k+1}}\le T_n f<\frac{1}{2^k}\}}(x)\Big)^{\frac{1}{2}} \\ &\le \Big(\sum_{k=0}^{\infty} \frac{1}{2^{2k}} \cdot \sum_{n=1}^{\infty} 1_{\{\frac{1}{2^{k+1}}\le T_n f<\frac{1}{2^k}\}}(x)\Big)^{\frac{1}{2}}. \end{aligned}$$

It follows again by Chebychev's inequality and the assumption of the theorem that

$$\begin{aligned} \mu(\{S_1 \ge 1\}) &= \mu(\{S_1^2 \ge 1\}) \\ &\le \mu\Big(\Big\{\sum_{k=0}^{\infty} \frac{1}{2^{2k}} \cdot \sum_{n=1}^{\infty} 1_{\{\frac{1}{2^{k+1}}\le T_n f<\frac{1}{2^k}\}} \ge 1\Big\}\Big) \\ &\le \sum_{k=0}^{\infty} \frac{1}{2^{2k}} \cdot \sum_{n=1}^{\infty} \mu(\{T_n f \ge \frac{1}{2^{k+1}}\}) \\ &\le \sum_{k=0}^{\infty} \frac{1}{2^{2k}} \cdot C \cdot 2^{k+1}\|f\|_1 \\ &= 4C\|f\|_1. \end{aligned}$$

https://doi.org/10.4153/CJM-1999-044-2 Published online by Cambridge University Press

Hence,

$$\begin{aligned}\mu(\{Sf \geq 2\}) &\leq \mu(\{S_1 f \geq 1\}) + \mu(\{S_2 f \geq 1\}) \\ &\leq 5C\|f\|_1.\end{aligned}$$

This proves the theorem. ■

Remark 2.4 a) Theorem 2.1 can be used together with Theorem 2.3 to give the previous result Theorem 2.2.

b) There is no assumption about the operators commuting with a mixing family of transformations. For these reasons, this argument will apply in cases where the Stein-Sawyer principle does not apply. But also, large deviation results can never be the only route to square functions inequalities. For instance, let $A_n f(x) = \frac{1}{n}\sum_{k=1}^n f(\tau^k x)$ be the usual average in ergodic theory. Then by the result in Jones, Ostrovskii and Rosenblatt [12], there is a weak inequality in $L_1(X)$ for the square function

$$Sf(x) = \Big(\sum_{n=1}^{\infty}\Big|A_{2^{n+1}} f(x) - A_{2^n} f(x)\Big|^2\Big)^{1/2}.$$

However, also in [12] it is shown that the operators $T_n = |A_{2^{n+1}} - A_{2^n}|$ do not satisfy the large deviation inequality in the assumption of Theorem 2.3.

The argument in the proof of Theorem 2.2 can be used to prove a result for functions more general than the square function. The precise result is as follows.

Theorem 2.5 *Assume that ϕ is a positive function defined on the positive real numbers which is increasing near zero. Assume $\sum_{n=1}^{\infty}\phi(\frac{1}{n}) < \infty$. For any sequence of moving averages $M_n f(x)$ with non-decreasing lengths (L_n), the series*

$$\sum_{n=1}^{\infty}\phi\Big(\Big|\frac{M_n f(x)}{n}\Big|\Big)$$

is finite a.e. x for any positive $f \in L_1(X)$.

Remark 2.6 Here if $\phi(x) = x^p$ for some $p > 1$, then actually the argument of Theorem 2.3 would work.

Proof As part of the large deviation result Theorem 2.1, we know that $|\frac{M_n f(x)}{n}|$ converges to 0 a.e. for any $f \in L_1(X)$. Hence, to prove the result, it suffices to consider only the series

$$\sum_{n=1}^{\infty}\phi\Big(\Big|\frac{M_n f(x)}{n}\Big|\Big)$$

excluding the terms n for which $|\frac{M_n f(x)}{n}| \geq 1$. For the same reason, there is no harm in assuming that ϕ is increasing on $[0,\infty)$.

Now, for a.e. x, we can take some $n(x) \geq 1$ such that for any $n \geq n(x)$, $|\frac{M_n f(x)}{n}| < 1$. Now as in Theorem 2.2,

$$\sum_{n=n(x)}^{\infty} \phi\Big(\Big|\frac{M_n f(x)}{n}\Big|\Big) = \sum_{n=n(x)}^{\infty} \phi\Big(\Big|\frac{M_n f(x)}{n}\Big|\Big) \sum_{m=0}^{\infty} 1_{\{\frac{1}{2^{m+1}} \leq |\frac{M_n f}{n}| < \frac{1}{2^m}\}}(x)$$
$$\leq \sum_{m=0}^{\infty} \sum_{n=n(x)}^{\infty} \phi\Big(\frac{1}{2^m}\Big) 1_{\{\frac{1}{2^{m+1}} \leq |\frac{M_n f}{n}| < \frac{1}{2^m}\}}(x).$$

Thus, by Theorem 2.1 once more,

$$\int \sum_{n=n(x)}^{\infty} \phi\Big(\Big|\frac{M_n f(x)}{n}\Big|\Big)\, d\mu(x) \leq \sum_{m=0}^{\infty} \sum_{n=1}^{\infty} \phi\Big(\frac{1}{2^m}\Big) \mu\Big(\Big\{\frac{1}{2^{m+1}} \leq \Big|\frac{M_n f}{n}\Big|\Big\}\Big)$$
$$\leq C \sum_{m=0}^{\infty} 2^{m+1} \phi\Big(\frac{1}{2^m}\Big) \|f\|_1.$$

The condensation principle for series is the elementary fact that if (a_n) is a decreasing sequence of positive real numbers, then the series $\sum_{n=1}^{\infty} a_n$ converges if and only if the condensed series $\sum_{n=1}^{\infty} 2^n a_{2^n}$ converges. This applies to the convergent series $\sum_{n=1}^{\infty} \phi(\frac{1}{n})$, showing that $\sum_{m=0}^{\infty} 2^{m+1} \phi(\frac{1}{2^m})$ is also convergent. This means that $\sum_{n=n(x)}^{\infty} \phi(|\frac{M_n f(x)}{n}|)$ is integrable and hence finite a.e. for any positive $f \in L_1(X)$. Hence, the entire series $\sum_{n=1}^{\infty} \phi(|\frac{M_n f(x)}{n}|)$ is convergent a.e. ∎

Under the type of assumptions on (T_n) as in Theorem 2.3, and with some additional assumptions about ϕ, we can actually get a result that generalizes Theorem 2.3 to include Theorem 2.5, strengthened to give a weak $(1,1)$ estimate too. Instead, let us just observe two similar versions of this type of theorem.

Corollary 2.7 *Let $p > 1$. For moving averages $M_n f(x)$ with non-decreasing lengths (L_n), the operator*

$$S_p f(x) = \Big(\sum_{n=1}^{\infty} \Big|\frac{M_n f(x)}{n}\Big|^p\Big)^{1/p}$$

is finite a.e for any $f \in L_1(X)$, and satisfies a weak type $(1,1)$ inequality on $L_1(X)$.

Proof The finiteness of the series a.e. is just Theorem 2.5. The weak estimate follows as in Theorem 2.2 or in Theorem 2.3 . ∎

In the particular case where the lengths (L_n) for the moving averages are all 1, Corollary 2.7 gives this next result, that also appears when $p = 2$ and $v_n = n$ in Jones [11]. Actually, the argument in Jones [11] can be seen to give this more general result too. The proof here is the same as the one just given for Corollary 2.7.

https://doi.org/10.4153/CJM-1999-044-2 Published online by Cambridge University Press

Corollary 2.8 *Let (v_n) be any sequence in $\mathbb{Z}$. Then for any $p > 1$,*

$$S_p f(x) = \left(\sum_{n=1}^{\infty} \left| \frac{f(\tau^{v_n} x)}{n} \right|^p \right)^{1/p}$$

is finite a.e for any $f \in L_1(X)$, and satisfies a weak type $(1,1)$ inequality on $L_1(X)$.

A standard argument with series which is even easier here because the terms are positive shows that we have the following corollary which is interesting even in the case that all $v_n = n$.

Corollary 2.9 *Let (v_n) be any sequence in $\mathbb{Z}$. Let $p > 1$ and let $f \in L_1(X)$. Then*

$$\lim_{n\to\infty} \frac{1}{n^p} \sum_{k=1}^{n} |f(\tau^{v_k} x)|^p = 0 \quad \textit{a.e.}$$

Remark 2.10 Assume that $v_n = n$ for all $n \geq 1$ in the above. Then the result is not immediate from the Pointwise Ergodic Theorem because just knowing that f is in $L_1(X)$ of course does not guarantee that $|f|^p$ is in $L_1(X)$. But this is also why one has to divide the sum by n^p and not just by n as in the Pointwise Ergodic Theorem.

Because any positive function $h \in L_1(X)$ is of the form $h = f^p$ for the positive function $f = h^{1/p} \in L_p(X) \subset L_1(X)$, Corollary 2.9 gives this result.

Corollary 2.11 *Let (v_n) be any sequence in $\mathbb{Z}$. Let $p > 1$ and let $h \in L_1(X)$. Then*

$$\lim_{n\to\infty} \frac{1}{n^p} \sum_{k=1}^{n} |h(\tau^{v_k} x)| = 0 \quad \textit{a.e.}$$

This is not a difficult fact to see by other means. Indeed, the series $\sum_{k=1}^{\infty} \frac{|h(\tau^{v_k} x)|}{k^p}$ is integrable for any $p > 1$ and $f \in L_1(X)$. Thus, the series also converges absolutely a.e. and Corollary 2.11 follows immediately. The more interesting question is what can be put in place of the factor $\frac{1}{n^p}$? Of course, generally the averages $\frac{1}{n} \sum_{k=1}^{n} |h(\tau^{v_k} x)|$ do not converge a.e., and so some extra factor is needed beyond just the factor n. The simple integration argument here shows that any factor $\frac{1}{L_n}$ such that $\sum_{n=1}^{\infty} \frac{1}{L_n} < \infty$ will work in place of $\frac{1}{n^p}$ with $p > 1$. But it is not clear what the correct factor to divide by is here. This issue is discussed further in Akcoglu, Jones, and Rosenblatt [1]. However, the next few results at least answer the question of whether the divisor n^p is the correct divisor in general in the context of Corollary 2.9.

Proposition 2.12 *Assume that τ is an ergodic transformation. If (a_n) is any sequence with $\lim_{n\to\infty} a_n = \infty$, and (v_n) is a sequence in $\mathbb{Z}$ such that $\sup_{n\geq 1} v_n = \infty$, then there exists a positive function $f \in L_1(X)$ such that*

$$\sup_n \frac{f(\tau^{v_n} x)}{n} a_n = \infty \quad \textit{a.e.}$$

Proof Assume first that the v_n are distinct. Assume there is a weak inequality on $\mathbb{Z}$ of the form, for any $\lambda > 0$,

$$\operatorname{card}\left\{k \in \mathbb{Z} : \sup_n \frac{a_n \phi(k + v_n)}{n} \geq \lambda\right\} \leq \frac{C}{\lambda} \|\phi\|_1.$$

Take $\phi = \delta_0$, the Dirac mass at 0. This weak inequality becomes

$$\operatorname{card}\left\{k \in \mathbb{Z} : k = -v_n \text{ for some } n \text{ with } \frac{a_n}{n} \geq \lambda\right\} \leq \frac{C}{\lambda}.$$

The left hand side of this weak inequality is then $\operatorname{card}\{n \geq 1 : \frac{a_n}{n} \geq \lambda\}$. Suppose though that $a_n \geq A$ for any $n \geq n_A$. Then for any $\lambda > 0$,

$$\begin{aligned}\operatorname{card}\left\{n \geq 1 : \frac{a_n}{n} \geq \lambda\right\} &\geq \operatorname{card}\left\{n \geq n_A : \frac{A}{n} \geq \lambda\right\} \\ &= \operatorname{card}\left\{n \geq n_A : \frac{A}{\lambda} \geq n\right\} \\ &\geq \frac{A}{\lambda} - n_A.\end{aligned}$$

Hence, $\frac{C}{\lambda} \geq \frac{A}{\lambda} - n_A$ for any $\lambda > 0$. This implies $A \leq C$. But A can be as large as we like. Hence, there cannot be a weak inequality.

When the (v_n) are not distinct, then the inequality becomes a similar one with an even larger value of a_n corresponding to the distinct values of v_n. So again there is no weak inequality. The Stein-Sawyer theorem, see Garsia [9], and the ergodicity of τ shows that failure of the weak inequality implies that there exists a positive function $f \in L_1(X)$ such that $\sup_n \frac{f(\tau^{v_n} x)}{n} a_n = \infty$ a.e. ■

The following is then an immediate consequence of Proposition 2.12.

Proposition 2.13 *Let (a_n) be any sequence with $\lim_{n\to\infty} a_n = \infty$. Then for each p, $1 < p < \infty$, there is a positive function $f \in L_1(X)$ such that the ratios $R_n f(x) = \frac{a_n}{n^p} \sum_{k=1}^n f^p(\tau^k x)$ diverge a.e.*

The result in Proposition 2.12 also shows that Corollary 2.8 cannot be improved in the sense that if (a_n) is any sequence tending to ∞, and τ is ergodic, then for any p, $1 < p < \infty$, there exists a positive function $f \in L_1(X)$ such that $\sum_{n=1}^{\infty} (\frac{f(\tau^{v_n} x)}{n})^p a_n = \infty$ a.e.

The conclusion is also that one cannot generally improve on the results in Corollary 2.7. One might think though, that it is possible to improve Corollary 2.7 if the lengths L_n were tending to infinity. The next result shows that even with this additional assumption on (L_n), one cannot improve Corollary 2.7.

Proposition 2.14 *Given any sequence (a_n) with $\lim_{n\to\infty} a_n = \infty$, there exists a sequence $(M_n f)$ of moving averages with non-decreasing lengths (L_n) such that $\lim_{n\to\infty} L_n = \infty$, but such that for a generic set of functions $f \in L_1(X)$,*

$$\left(\sum_{n=1}^{\infty} \left|\frac{a_n M_n f(x)}{n}\right|^p\right)^{1/p} = \infty \quad a.e.$$

Proof This is immediate from the result in [16] which gives the same divergence for the term $\frac{a_n M_n f(x)}{n}$. ■

However, there is another way to describe the rate result in Corollary 2.9 that is successful. We state this just for the case $p = 2$ to illustrate the point. The proof follows easily from Corollary 2.8 by partial summation.

Proposition 2.15 *For any positive function $f \in L_1(X)$, the series $\sum_{n=1}^{\infty} \left(\frac{1}{n^2} \sum_{k=1}^{n} f^2(\tau^k x)\right) \frac{1}{n}$ converges a.e.*

3 Counting Jump Oscillations

Inherent in the proofs of the main results in Section 2 are counting criteria that measure how often sums associated with ergodic transformations are in certain intervals of the range. In the same sense, upcrossing and jump inequalities have been studied in ergodic theory to get a better understanding of the behavior of the classical ergodic averages. This was originally done in Bishop [4]. Recently, these issues were considered in Kalikow and Weiss [15] and in Jones, Kaufman, Rosenblatt, and Wierdl [13]. The jump inequality obtained in [13] is discussed in this section and it is shown that it is the best possible result for jumps, although one can do better when dealing just with upcrossings, in both the case of martingales and ergodic averages.

3.1 The Martingale Case

Let $f = (f_n)$ be a martingale, and let $\Lambda(\lambda, f, x)$ denote the number of λ jumps. That is, let $\Lambda(\lambda, f, x) = \max\{n :$ there exist $s_1 < t_1 \leq s_2 < t_2 \cdots \leq s_n < t_n$, such that $|f_{t_k}(x) - f_{s_k}(x)| \geq \lambda\}$.

It is well-known that

$$\mu\big(\{x : \Lambda(\lambda, f, x) \geq n\}\big) \leq \frac{C}{\lambda\sqrt{n}} \|f\|_1$$

where as usual $\|f\|_1 = \sup_{n\geq 1} \|f_n\|_1$. See for example Kachurovskii [14], Theorem 35, or Jones, Kaufman, Rosenblatt, and Wierdl [13], Section 6. Also, from either of these articles one can see this more general result.

Theorem 3.1 *For $1 \leq p < \infty$ we have*

$$\mu\big(\{x : \Lambda(\lambda, f, x) > n\}\big) \leq \frac{c}{\lambda^p n^{p/2}} \|f\|_p^p.$$

Remark 3.2 a) The article Kachurovskii [14] is really worth close examination because it contains a number of interesting theorems; in particular it contains results and questions related to the article by Jones, Kaufman, Rosenblatt, and Wierdl [13] as well as to this present article.

b) The constant here in Theorem 3.1 does not depend on p.

Rewriting this result for $p = 1$, we see that

$$\lambda\sqrt{n}\mu\big(\{x : \Lambda(\lambda, f, x) \geq n\}\big) \leq C\|f\|_1. \tag{3.1}$$

For upcrossings, there is a similar inequality. For any $\alpha < \beta$, let $N(\alpha, \beta, f, x)$ denote the number of upcrossings between α and β. That is, we take $N(\alpha, \beta, f, x)$ to be the maximum value of n such that there exists $s_1 < t_1 \leq s_2 < t_2 \cdots \leq s_n < t_n$ for which $f_{t_k} \leq \alpha$ and $f_{s_k} \geq \beta\}$. Then it has been shown that

$$(\beta - \alpha)n\mu\big(\{x : N(\beta, \alpha, f, x) \geq n\}\big) \leq C\|f\|_1. \tag{3.2}$$

This can be found in Doob [8] or in Bishop [4, 5].

It is natural to ask if the inequality for jumps can be improved, and in particular, can $\sqrt{n}$ in inequality (3.1) be improved to obtain n as in (3.2). In fact we will show that Theorem 3.1 is a sharp result by proving the following.

Theorem 3.3 *Let $\Phi(n)$ denote any strictly increasing function with $\lim_{n\to\infty}\Phi(n) = \infty$. Then for each p, $1 \leq p < \infty$, there is a martingale $f = (f_1, f_2, \dots)$ such that $f \in L^p$, but such that we have*

$$\sup_n n^{p/2}\Phi(n)\mu\big(\{x : \Lambda(1, f, x) > n\}\big) = \infty. \tag{3.3}$$

Remark 3.4 Take for instance $p = 1$ and $\Phi(n) = \sqrt{n}$ to see that the analog of (3.2) does not hold for jumps. Indeed, nothing faster than $\sqrt{n}$ by itself will work on $L_1(X)$.

Proof For a given pair of increasing sequences (N_k) and (n_k) of positive integers, we define a dyadic martingale f as follows. Let $s_1 = 0$, and for $k > 1$, define s_k to be the maximum of $\{s_{k-1} + N_{k-1}, n_k\}$. Define $d_k = \sum_{j=1}^{N_k} r_{j+s_k}$ where as usual, r_k denotes the kth Rademacher function. Note that since $s_k \geq n_k$, all d_k will be periodic with period $\frac{1}{2^{n_k}}$.

Define

$$f(t) = \sum_{k=1}^{\infty} d_k(t)1_{[0,1/2^{n_k})}(t).$$

Then, using the fact that the L_p-norms of Rademacher sums are all equivalent, we get

https://doi.org/10.4153/CJM-1999-044-2 Published online by Cambridge University Press

for a suitable constant c_p,

$$\begin{aligned}\|f\|_p &= \Big(\int_0^1 \Big|\sum_{k=1}^{\infty} d_k(t)1_{[0,1/2^{n_k})}(t)\Big|^p \,d\mu(t)\Big)^{\frac{1}{p}}\\ &\le \sum_{k=1}^{\infty}\Big(\int_0^1 |\,d_k(t)1_{[0,1/2^{n_k})}(t)|^p \,d\mu(t)\Big)^{\frac{1}{p}}\\ &\le \sum_{k=1}^{\infty}\Big(\frac{1}{2^{n_k}}\int_0^1 |\,d_k(t)|^p \,d\mu(t)\Big)^{\frac{1}{p}}\\ &\le \sum_{k=1}^{\infty}\Big(\frac{1}{2^{n_k}}\Big)^{\frac{1}{p}}\|d_k\|_p\\ &\le c_p\sum_{k=1}^{\infty}\Big(\frac{1}{2^{n_k}}\Big)^{\frac{1}{p}}\|d_k\|_2\\ &\le c_p\sum_{k=1}^{\infty}\Big(\frac{1}{2^{n_k}}\Big)^{\frac{1}{p}}\sqrt{N_k}.\end{aligned}$$

Fix p, $1 \le p < \infty$. For each integer $k > 0$ select N_k so that $\Phi(N_k) \ge k2^k$. Now generally

$$\mu\big(\{x : \Lambda(1, f, x) > N_k\}\big) \ge \mu\big(\{x : \Lambda(1, d_k 1_{[0,1/2^{n_k})}, x) > N_k\}\big) \ge \frac{1}{2^{n_k}}.$$

Consequently, if we take n_k such that we have $N_k^{p/2}2^k \le 2^{n_k} \le 2N_k^{p/2}2^k$, then we have

$$N_k^{p/2}\Phi(N_k)\mu\big(\{x : \Lambda(1, f, x) > N_k\}\big) \ge N_k^{p/2}k2^k\frac{1}{2^{n_k}} \ge \frac{k}{2}.$$

Hence if we could show that $f \in L^p$, we would be done. However, using the estimate above, and since our choice of n_k implies $\sqrt{N_k} \le (2^{n_k}/2^k)^{1/p}$, we see that

$$\|f\|_p \le c_p\sum_{k=1}^{\infty}\Big(\frac{1}{2^{n_k}}\Big)^{\frac{1}{p}}\Big(\frac{2^{n_k}}{2^k}\Big)^{\frac{1}{p}} \le c_p\sum_{k=1}^{\infty}\frac{1}{2^{k/p}} < \infty. \qquad \blacksquare$$

Remark 3.5 a) The argument did not depend on the fact that we used the Rademacher functions in the natural order. We could start at any Rademacher function we like, and as long as all the Rademacher functions are different, we could use any subsequence of them to complete the construction. All we really used was that the Rademacher functions are independent, and take only the values ± 1. By repeating the argument with later Rademacher functions at each stage, and adding the results with suitable normalizations, we can create a bounded function f such that

$$\sup_{\lambda,n}\lambda\sqrt{n}\phi(n)\mu\big(\{x : \Lambda(\lambda, f, x) \ge n\}\big) = \infty. \tag{3.4}$$

To get this unboundedness for bounded functions does require though that one vary the scale factor λ, otherwise it would be contradicting Theorem 3.1.

b) Theorem 3.1 also shows that the L^p function constructed in Theorem 3.3 cannot be in any L^r, for $r > p$. Hence neither Theorem 3.3 nor Theorem 3.1 can be improved for any finite p.

3.2 The Integer Case

We want to establish the above result in the ergodic theory setting, but before we do that we need to establish the analogous result on $\mathbb{Z}$. On $\mathbb{Z}$ we will consider the analog of the Rademacher functions. For each $n \geq 0$ we define $\phi_n : \mathbb{Z} \to \{\pm 1\}$ by $\phi_n(k) = 1$ if $0 \leq k < 2^n$, $\phi_n(k) = -1$ if $2^n \leq k < 2^{n+1}$, and ϕ_n is periodic with period 2^{n+1}.

We also use the "density measure" D, defined as follows. Let ψ be a function from $\mathbb{Z}$ to $\mathbb{R}$. Let

$$\bar{D}(\psi) = \limsup_{L \to \infty} \frac{1}{2L+1} \sum_{x=-L}^{L} \psi(x)$$

and

$$\underline{D}(\psi) = \liminf_{L \to \infty} \frac{1}{2L+1} \sum_{x=-L}^{L} \psi(x).$$

Let $D(\psi) = \underline{D}(\psi)$ when $\underline{D}(\psi) = \bar{D}(\psi)$. While $D(\psi)$ may fail to be defined, $\bar{D}(\psi)$ and $\underline{D}(\psi)$ will always exist, although they may not be finite. For a set B, let $D(B)$, $\bar{D}(B)$ and $\underline{D}(B)$ denote $D(1_B)$, $\bar{D}(1_B)$ and $\underline{D}(1_B)$ respectively. Also, for $1 \leq p < \infty$, let $\|\psi\|_p$ denote $\big(\lim_{L\to\infty} \frac{1}{2L+1} \sum_{x=-L}^{L} |\psi(x)|^p\big)^{1/p}$ when this limit exists.

With this notation, we have the following version of the Weak Law of Large Numbers.

Proposition 3.6 *Let $\psi_1, \psi_2, \ldots$ satisfy the following:*

1. *There is a finite constant μ such that $D(\psi_i)$ exists and $D(\psi_i) = \mu$ for each i.*
2. *There is a finite constant σ^2 such that $D([\psi_i - \mu]^2)$ exists and $D([\psi_i - \mu]^2) = \sigma^2$ for each i.*
3. *For all $i \neq j$ we have $D(\psi_i \psi_j)$ exists and $D\big((\psi_i - \mu)(\psi_j - \mu)\big) = 0$.*

Then $\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \psi_i(x) = \mu$ in the sense that for each $\epsilon > 0$,

$$\lim_{N\to\infty} \bar{D}\Big(\{x : \Big|\frac{1}{N}\sum_{i=1}^{N} \psi_i(x) - \mu\Big| > \epsilon\}\Big) = 0.$$

https://doi.org/10.4153/CJM-1999-044-2 Published online by Cambridge University Press

Proof For fixed $N > 0$ we have

$$\begin{aligned}
\bar{D}\Big\{\Big|\frac{1}{N}\sum_{i=1}^{N}\psi_i-\mu\Big|>\epsilon\Big\} &= \limsup_{L\to\infty}\frac{1}{2L+1}\sum_{x=-L}^{L}1_{\{|\frac{1}{N}\sum_{i=1}^{N}\psi_i-\mu|>\epsilon\}}(x)\\
&\le \limsup_{L\to\infty}\frac{1}{2L+1}\sum_{x=-L}^{L}\frac{1}{\epsilon^2}\Big(\frac{1}{N}\sum_{i=1}^{N}\psi_i(x)-\mu\Big)^2\\
&= \frac{1}{\epsilon^2}\frac{1}{N^2}\limsup_{L\to\infty}\frac{1}{2L+1}\sum_{x=-L}^{L}\sum_{i=1}^{N}\sum_{j=1}^{N}\big(\psi_i(x)-\mu\big)\big(\psi_j(x)-\mu\big)\\
&\le \frac{1}{\epsilon^2}\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\limsup_{L\to\infty}\frac{1}{2L+1}\sum_{x=-L}^{L}\big(\psi_i(x)-\mu\big)\big(\psi_j(x)-\mu\big)\\
&= \frac{1}{\epsilon^2}\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}D\big((\psi_i-\mu)(\psi_j-\mu)\big)\\
&\le \frac{1}{\epsilon^2}\frac{1}{N^2}\sum_{i=1}^{N}\sigma^2\\
&= \frac{\sigma^2}{\epsilon^2}\frac{1}{N}.
\end{aligned}$$

Since ϵ and σ^2 are fixed, we see that the limit as N goes to ∞ will be zero. ■

We can now use Proposition 3.6 to derive a result similar to Theorem 3.3 for for functions on $\mathbb{Z}$. We need here the analogue of $\Lambda(\lambda, f, x)$. Let $A_n\psi(j) = \frac{1}{n}\sum_{k=1}^{n}\psi(j+k)$. Let $\Lambda(\lambda,\psi,j)$ denote the number of λ jumps of the averages $A_n\psi(j)$. That is, let $\Lambda(\lambda,\psi,j) = \max\{n : \text{there exist } s_1 < t_1 \le s_2 < t_2 \cdots \le s_n < t_n, \text{ such that } |A_{t_k}\psi(j) - A_{s_k}\psi(j)| \ge \lambda\}$.

Theorem 3.7 *If $\Phi(n)$ denotes any strictly increasing function with $\lim_{n\to\infty}\Phi(n) = \infty$, then for each p, $1 \le p < \infty$, and each $\beta > 0$, there is a function ϕ with $\|\phi\|_p \le 1$ such that we have*

$$\sup_n n^{p/2}\Phi(n)D\big(\{j : \Lambda(1,\phi,j)\ge n\}\big) > \beta. \tag{3.5}$$

Proof Fix d and L, positive integers to be determined later, and let $\Delta_\ell\phi(j) = A_{2^{d\ell}}\phi(j) - A_{2^{d\ell-1}}\phi(j)$. We will make some preliminary observations concerning the averages of the functions ϕ_n introduced above.

First, for $m < k$, $\Delta_k\phi_{dm}(j) = 0$ for all j, since we are taking long averages compared to the period of ϕ_{dm}, each average covers an integer multiple of a full period of ϕ_{dm}, and ϕ_{dm} has mean value zero. Also a simple computation shows that $|\Delta_\ell\phi_{d\ell}| \ge \frac{1}{2}$ on a set B_ℓ, with B_ℓ periodic, and $D(B_\ell) \ge \frac{1}{4}$. We can take B_ℓ to be the second quarter of each dyadic interval of length $2^{d\ell}$. Here $D(B_\ell)$ exists because B_ell is periodic. But also for $m > \ell$, $\Delta_\ell\phi_{dm}(j)$ can be as large as 1, yet can be non-zero only at the last $2^{d\ell}$ points at the ends of dyadic intervals of length 2^{dm}. Hence if $E(\ell, m)$ denotes the set where $\Delta_\ell \phi_{dm}(j) \neq 0$, then $D\big(E(\ell, m)\big) \leq 2^{d(\ell-m)}$.

The sets B_ℓ are independent in the sense that

$$D(B_\ell \cap B_j) = D(B_\ell)D(B_j),$$

so the functions $1_{B_\ell} - D(B_\ell)$ are orthogonal with respect to D. Thus, Proposition 3.6 implies that if N is large enough, then $\frac{1}{N}\sum_{k=1}^{N} 1_{B_\ell} > \frac{1}{8}$ on a set B, with $D(B) > \frac{7}{8}$. Let $E_\ell = \bigcup_{m=\ell+1}^{L} E(\ell, m)$. Then

$$\begin{aligned} D(E_\ell) &\leq \sum_{m=\ell+1}^{L} D\big(E(\ell, m)\big) \\ &\leq \sum_{m=\ell+1}^{\infty} \frac{1}{2^{d(m-\ell)}} \\ &\leq \frac{2}{2^d}. \end{aligned}$$

We are now ready to define our function ϕ. As in the proof of the martingale case, we will need two increasing sequences, (N_k) and (n_k). Define $s_1 = 0$ and select $s_k \geq s_{k-1} + dN_{k-1}$. We will impose additional conditions on (s_k) later, but for now note that it can grow as fast as we like. Let $d_k = \sum_{j=1}^{N_k} \phi_{s_k+dj}$. Let $\chi_k = 1_{[0, 8d \times 2^{dN_k+s_k})}$ on $[0, n_k \times d \times 2^{dN_k+s_k})$ and extend χ_k to be periodic on $\mathbb{Z}$. Here we can assume without loss of generality that $n_k > 8$ for all k. Define $\phi = \sum_{k=1}^{L} d_k \chi_k$.

Since ϕ is periodic, $\|\phi\|_p$ is well-defined. As before, we need to check

$$\begin{aligned} \|\phi\|_p &= \big(D(|\phi|^p)\big)^{\frac{1}{p}} \\ &= \Big(\lim_{R\to\infty} \frac{1}{2R+1} \sum_{j=-R}^{R} |\phi(j)|^p\Big)^{\frac{1}{p}} \\ &= \lim_{R\to\infty} \Big(\frac{1}{2R+1} \sum_{j=-R}^{R} |\sum_{k=1}^{L} d_k(j)\chi_k(j)|^p\Big)^{\frac{1}{p}} \\ &\leq \lim_{R\to\infty} \sum_{k=1}^{L} \Big(\frac{1}{2R+1} \sum_{j=-R}^{R} |d_k(j)\chi_k(j)|^p\Big)^{\frac{1}{p}} \\ &\leq \sum_{k=1}^{L} \Big(\frac{8}{n_k}\Big)^{\frac{1}{p}} \lim_{R\to\infty} \Big(\frac{1}{2R+1} \sum_{j=-R}^{R} |d_k(j)|^p\Big)^{\frac{1}{p}} \\ &\leq c_p \sum_{k=1}^{L} \Big(\frac{1}{n_k}\Big)^{\frac{1}{p}} \lim_{R\to\infty} \Big(\frac{1}{2R+1} \sum_{j=-R}^{R} |d_k(j)|^2\Big)^{\frac{1}{2}} \\ &\leq c_p \sum_{k=1}^{L} \Big(\frac{1}{n_k}\Big)^{\frac{1}{p}} \sqrt{N_k}, \end{aligned}$$

where in the next to the last step we used a version of Khinchine's inequality, and in the final step, the orthogonality of the sequence (ϕ_j).

Next select N_k so that $\Phi(N_k) > k2^{kp}/c_p^p$, and then select n_k such that $N_k^{p/2}2^{kp} \le c_p^p n_k \le 2 \times N_k^{p/2}2^{kp}$. The remainder of the proof involve making estimates of D on various periodic sets obtained from ϕ, and so at least the value of D will always be well-defined. First, we want to show that

$$D\Big(\Big\{ j : \Lambda\Big(\frac{1}{4}, \phi, j\Big) > \frac{N_k}{16}\Big\}\Big) \ge D\Big(\Big\{ j : \Lambda\Big(\frac{1}{2}, d_k\chi_k, j\Big) > \frac{N_k}{16}\Big\}\Big) \ge \frac{1}{n_k}$$

Then we will have as before

$$N_k^{p/2}\Phi(N_k)D\Big(\Big\{ j : \Lambda\Big(\frac{1}{4}, \phi, j\Big) > N_k\Big\}\Big) \ge N_k^{p/2}(k2^{kp}/c_p^p)\frac{1}{n_k} \ge \frac{k}{2}.$$

Taking $L > 2\beta$ will allow us to obtain the conclusion of inequality (3.5). Our choices of (N_k) and (n_k), and the computation of $\|\phi\|_p$, allow us to conclude that $\|\phi\|_p \le 1$. Hence we just need to obtain the desired estimate for $D(\{ j : \Lambda(\frac{1}{4}, \phi, j) > \frac{N_k}{16}\})$.

For $s_k \le \ell \le s_k + N_k$, we have $|\Delta_\ell(d_k)(j)| \ge \frac{1}{2}1_{B_k\setminus E_k}(j)$, since $|\Delta_\ell(\phi_{d\ell})(j)| \ge \frac{1}{2}$ on B_ℓ, and $\Delta_\ell(\phi_{dm})(j) = 0$ if $k \ne m$ and $j \notin E_k$. Therefore,

$$\begin{aligned}\Lambda\Big(\frac{1}{2}, d_k, j\Big) &\ge \sum_{\ell=s_k}^{s_k+N_k} 1_{B_\ell\setminus E_\ell}(j)\\ &\ge \sum_{\ell=s_k}^{s_k+N_k} 1_{B_\ell}(j) - \sum_{\ell=s_k}^{s_k+N_k} 1_{E_\ell}(j)\\ &= G(j) - H(j).\end{aligned}$$

From Proposition 3.6, we know that if N_k is chosen large enough, $G(j)$ is at least $\frac{1}{2}\frac{N_k}{4}$ on the set B, with $D(B) \ge \frac{7}{8}$. If we could show that $H(j)$ is at most $\frac{N_k}{16}$ on a set C, with $D(C) > \frac{7}{8}$, then we would have $\Lambda(\frac{1}{2}, d_k, \cdot) > \frac{N_k}{16}$ on the set $B \cap C$, with $D(B \cap C) > \frac{3}{4}$.

We have

$$\begin{aligned}D\Big(\Big\{ j : H(j) > \frac{N_k}{16}\Big\}\Big) &\le \frac{16}{N_k}D(H)\\ &\le \frac{16}{N_k}D\Big(\sum_{\ell=s_k}^{s_k+N_k} 1_{E_\ell}\Big)\\ &\le \frac{16}{N_k}N_k \times \frac{2}{2^d}\\ &\le \frac{32}{2^d}.\end{aligned}$$

Taking $d = 8$ shows $D(\{ j : H(j) > \frac{N_k}{16}\}) \le \frac{1}{8}$.

Let $C = \{j : H(j) \leq \frac{N_k}{16}\}$. If $j \in B \cap C$, then $G(j) - H(j) \geq \frac{N_k}{8} - \frac{N_k}{16} = \frac{N_k}{16}$, and we know $D(B \cap C) \geq \frac{3}{4}$. Hence $\Lambda(\frac{1}{2}, d_k, \cdot) > \frac{N_k}{16}$ on the set $B \cap C$, with $D(B \cap C) > \frac{3}{4}$, which implies $D\big(\{\Lambda(\frac{1}{2}, d_k, j) > \frac{N_k}{16}\}\big) \geq \frac{3}{4}$.

We now need to use our result about d_k to imply a result about ϕ. This is not difficult. First note that for any ℓ, $s_k \leq \ell \leq s_k + N_k$, we have $\Delta_\ell\big(d_k(j)\chi_k(j)\big) = \Delta_\ell d_k(j)$ if j is in the second $2^{dN_k+s_k}$ steps of a period of χ_k. This follows because with ℓ in the given range, if we start in the second block of length $2^{dN_k+s_k}$, the operator Δ_ℓ only averages over points which are in the support of χ_k. Hence $D\big(\Lambda(\frac{1}{2}, d_k\chi_k, \cdot) \geq \frac{N_k}{16}\big) \geq \frac{1}{n_k}$.

Finally, if we select (s_k) increasing rapidly enough, and if $m \neq k$, then for $s_k \leq \ell \leq s_k+N_k$ we will have $\Delta_\ell d_m(j)$ as close to zero as we want, and in particular, less than $\frac{1}{4}$, for j in the second $2^{dN_k+s_k}$ steps of a period of χ_k. To see this, observe that the length of the blocks we are averaging over are either much shorter than the period associated with d_k, or so much longer that we average over a full period, and get zero. Hence, since for such j we have $\Delta_\ell\phi(j) = \sum_{k=1}^{L} \Delta_\ell d_k(j)\chi_k(j)$. We now use the above estimates, and the proof is complete. ■

3.3 The Ergodic Case

We can now state the same result for the ergodic case. Let (X, Σ, μ) denote a probability space, and $\tau \colon X \to X$ a measurable measure preserving ergodic transformation. For $f \colon X \to X$, define the operators $A_k f(x) = \frac{1}{k}\sum_{n=0}^{k-1} f(\tau^n x)$. As in the martingale case, define $\Lambda(\lambda, f, x)$ to be the number of λ jumps taken by the process $A_k f(x)$. That is, $\Lambda(\lambda, f, x) = \max\{n$: there exist $s_1 < t_1 \leq s_2 < t_2 \cdots \leq s_n < t_n$, such that $|A_{t_k} f(x) - A_{s_k} f(x)| \geq \lambda\}$.

We first state the following proposition which is almost immediate from the proof of Theorem 3.7.

Proposition 3.8 *If $\Phi(n)$ denotes any strictly increasing function with $\lim_{n\to\infty} \Phi(n) = \infty$, then for each p, $1 \leq p < \infty$, any $\epsilon > 0$, and any constant K, there is a function f such that $\|f\|_p \leq \epsilon$ such that we have*

$$\sup_n n^{p/2}\Phi(n)\mu\big(\{x : \Lambda(1, f, x) \geq n\}\big) \geq K.$$

Proof Initially, choose k sufficiently large so that $c_p(\frac{1}{2}\frac{1}{2^k})^{1/p} \leq \epsilon$ and such that $k > 2(4^p)K$. Here k is chosen so that with $\lambda = \frac{1}{4}$, we have $\frac{k}{2}\lambda^p > K$. Also, consider the function $d_k\chi_k$ constructed in the proof of Theorem 3.7. Let p_k denote the period of χ_k. Form a Rokhlin tower of height p_k, and with error less than $\frac{1}{p_k}$. Copy the function $d_k\chi_k$ to the tower, and define f to be this function on the tower, and zero off the tower. The condition that $c_p(\frac{1}{2}\frac{1}{2^k})^{1/p} \leq \epsilon$ guarantees that $\|f\|_p \leq \epsilon$. Hence, the desired conclusion follows from the properties of the function $d_k\chi_k$ and the choice that $k > 2(4^p)K$ once k is sufficiently large to allow for the effect of the error $\frac{1}{p_k}$. ■

This result certainly denies the possibility of there being an L_p-norm inequality that bounds $\sup_n \lambda^p n^{p/2}\Phi(n)\mu\big(\{x : \Lambda(\lambda, f, x) \geq n\}\big)$. But with a little more work we can construct a single function in the spirit of Theorem 3.7.

Theorem 3.9 *If $\Phi(n)$ denotes any strictly increasing function with $\lim_{n\to\infty}\Phi(n)=\infty$, then for each p, $1\le p<\infty$, there is a function f with $\|f\|_p<\infty$, such that we have*

$$\sup_n n^{p/2}\Phi(n)\mu\bigl(\{x:\Lambda(1,f,x)\ge n\}\bigr)=\infty.$$

This result is typical of a corollary of a result like the one in Proposition 3.8 that often needs to be proved. It is like what one gets when applying the method of Sawyer to an unbounded maximal function. However, here the quantities in question are not quite subadditive. For this reason, it is perhaps worthwhile to formulate a general principle which gives Theorem 3.9. Suppose that $G(\lambda,n)\colon L_p(X)\to\mathbb{R}$ for each value of λ and n. Here we take $\lambda>0$ and $n>0$. We say $G(\lambda,n)$ is continuous on $L_p(X)$ when $G(\lambda,n)f$ tends to 0 as $\|f\|_p$ tends to 0. We say that $\{G(\lambda,n):\lambda>0,n>0\}$ is quasi-subadditive if 1) for any $f\in L_p(X)$, any λ and any n, $G(\lambda,n)(-f)=G(\lambda,n)f$, and 2) there is a constant $C>0$, such that for any $f_1,f_2\in L_p(X)$, any λ and any n, we have $G(2\lambda,2n)(f_1+f_2)\le C\bigl(G(\lambda,n)f_1+G(\lambda,n)f_2\bigr)$.

Proposition 3.10 *Suppose that $\{G(\lambda,n):\lambda>0,n>0\}$ is a family of continuous functions on $L_p(X)$ for some p, $1\le p\le\infty$, which is quasi-subadditive. Suppose that for some fixed λ_o, for any $K,\epsilon>0$, there exists $f\in L_p(X)$ such that $\|f\|_p<\epsilon$ and $\sup_{n>0}G(\lambda_o,n)f\ge K$. Then there exists $\gamma_o>0$ and $f\in L_p(X)$ such that $\sup_{n>0}G(\gamma_o,n)f=\infty$.*

Proof We will choose a sequence (f_m) in $L_p(X)$ such that $\|f_m\|_p\le\epsilon_m$ where each $\epsilon_m>0$ and $\sum_{m=1}^\infty\epsilon_m<\infty$. Then the series $\sum_{m=1}^\infty f_m$ converges in $L_p(X)$. The choice of the functions is made inductively. Fix a sequence (K_m) such that $\lim_{m\to\infty}K_m=\infty$. The conditions needed on (ϵ_m) and (K_m) for this construction to work are also described inductively.

Let $\epsilon_1>0$ be arbitrary and choose f_1 such that $\|f_1\|_p\le\epsilon_1$ and for some n_1, we have $G(\lambda_o,n_1)f_1\ge K_1$. Assume now that we have chosen f_m,ϵ_m, and n_m, $m=1,\dots,M$, such that $\|f_m\|_p\le\epsilon_m\le\frac{1}{2^m}$ and $G(\lambda_o,n_m)f_m\ge K_m$, for all $m=1,\dots,M$. Let $S_M=\sum_{m=1}^M f_m$ with $S_0=0$. It is possible that we have $\sup_{n>0}G(\frac{\lambda_o}{4},\frac{n}{4})S_M=\infty$. If so, with $\gamma_o=\frac{\lambda_o}{4}$, there is nothing left to prove. If not, let $C_M=\sup_{n>0}G(\frac{\lambda_o}{4},n)S_M$ and continue with the inductive construction. Let $f=\sum_{m=1}^\infty f_m$. Also, let $T_{M+1}=\sum_{m=M+1}^\infty f_m$.

Now the quasi-subadditivity of the $G(\lambda,n)$ shows that for any λ and n, if $M\ge1$,

$$\begin{aligned}G(4\lambda,4n)f_M&\le C\bigl(G(2\lambda,2n)f+G(2\lambda,2n)(S_{M-1}+T_{M+1})\bigr)\\&\le C\,G(2\lambda,2n)f+C^2\bigl(G(\lambda,n)S_{M-1}+G(\lambda,n)T_{M+1}\bigr).\end{aligned}$$

By choosing the values of (ϵ_m) appropriately, the continuity hypothesis guarantees that we can arrange for $G(\frac{\lambda_o}{4},\frac{n_M}{4})T_{M+1}\le1$. Hence, we have for any $M\ge2$,

$$G\Bigl(\frac{\lambda_o}{2},\frac{n_M}{2}\Bigr)f\ge\frac{G(\lambda_o,n_M)f_M}{C}-C\,C_{M-1}-C.$$

Hence, for any $M\ge2$, we have

$$G\Bigl(\frac{\lambda_o}{2},\frac{n_M}{2}\Bigr)f\ge\frac{K_M}{C}-C\,C_{M-1}-C.$$

Therefore, with an appropriate inductive choice of (K_M), we can guarantee that $\sup_{n>0} G(\gamma_o, n)f = \infty$ with $\gamma_o = \frac{\lambda_o}{2}$. ■

Proof of 3.9 By replacing $\Phi(n)$ by a more slowly growing increasing function, we can assume without loss of generality that $\Phi(2n) \le 2\Phi(n)$. It is also not hard to see that $G(\lambda, n) = \mu\bigl(\{x : \Lambda(\lambda, f, x) \ge n\}\bigr)$ is quasi-subadditive with the constant $C = 1$. Therefore, we have the function $G(\lambda, n) = \lambda^p n^{p/2}\Phi(n)\mu\bigl(\{x : \Lambda(\lambda, f, x) \ge n\}\bigr)$ is quasi-subadditive. Thus, Proposition 3.8 and Proposition 3.10 complete the proof. ■

There is another interesting fact about the jump function that the method of this section can give us. First, recall that the upcrossing function $N(\alpha, \beta, f, x)$ actually satisfies a type of strong $(1, 1)$ estimate. In Bishop [4], it is shown that for any $\beta > \alpha$, we have $\int N(\alpha, \beta, f, x)\, d\mu(x) \le \frac{1}{\beta-\alpha} \int \bigl(f(x) - \alpha\bigr)^+ d\mu(x)$. This is a remarkable inequality because it so much like a strong $(1, 1)$ inequality. The same type of inequality holds for the upcrossing function for martingales. We have seen here that the correct gauge for jumps of martingales or ergodic averages on the other hand is $\sqrt{\Lambda(\lambda, f, x)}$. This function is weak $(1, 1)$ and strong (p, p) for all p, $1 < p < \infty$.

These remarks derive from similar results for square functions, so it is not unreasonable to ask if something better, like integrability, holds for $\sqrt{\Lambda(\lambda, f, x)}$ just as it did for $N(\alpha, \beta, f, x)$. This is the reason that the following three results are quite interesting. We thank Don Burkholder for allowing us to include his proof of Theorem 3.11. The same theorem is given in Kachurovskii [14], Theorem 34; see also Bourgain [6], inequality (3.5). Using the next martingale result will allow us to obtain the same result in the ergodic setting.

Theorem 3.11 *There does not exist a constant C such that for all martingales $f = (f_n)$, we have $\int \sqrt{\Lambda(1, f, x)}\, d\mu(x) \le C\|f\|_1$.*

Proof Let (r_k) be the Rademacher sequence. Let $f_n = \sum_{k=1}^n 1_{\{s\ge k\}} r_k$ where s is the stopping time $s(x) = \inf\{m \ge 1 : \sum_{k=1}^m r_k(x) = -1\}$. Then (f_n) is a martingale with $\|f\|_1 = \sup_{n\ge 1} \|f_n\|_1 < \infty$. Let $Sf = (\sum_{n=1}^\infty |f_n - f_{n+1}|^2)^{1/2}$ be the square function and $S_N f = (\sum_{n=1}^N |f_n - f_{n+1}|^2)^{1/2}$ be its partial sums. We have

$$\int_0^1 \sqrt{\Lambda(1, f_n, x)}\, dx = \int_0^1 \sqrt{\min\bigl(s(x), n\bigr)}\, dx$$
$$= \int_0^1 S_n f(x)\, dx.$$

But $\lim_{n\to\infty} \int_0^1 S_n f(x)\, dx = \int_0^1 S(x)\, dx$, which we claim is infinite, thus showing there can be no inequality as above. Indeed, assume that $\int_0^1 S(x)\, dx$ is finite. Let $f^\star = \sup_{n\ge 1} |f_n|$. Then $\int_0^1 f^\star\, dx$ would also be finite, since for the type of martingale we are considering, Burkholder and Gundy [7] have shown that for some constant C, $\|f^\star\|_1 \le C\|Sf\|_1$. Note that $\int f_n(x)\, dx = 0$ for every n. So because $f^\star \in L_1(X)$, we can apply the Dominated Convergence Theorem to conclude that

$$0 = \lim_{n\to\infty} \int f_n(x)\, dx = \int \lim_{n\to\infty} f_n(x)\, dx.$$

https://doi.org/10.4153/CJM-1999-044-2 Published online by Cambridge University Press

Since $\lim_{n\to\infty} f_n(x) = -1$ a.e., we see that $0 = -1$, which is a contradiction. ■

By taking the martingale $(f_1, \ldots, f_N)$, the initial segment of the martingale in Theorem 3.11, one can also see of course that there is no homogeneous inequality for the class of closed martingales. But more specifically, it is also easy to see this result.

Proposition 3.12 *There is a function $f \in L_1(X)$, some $\lambda > 0$, and an increasing sequence of σ-algebras (F_n) such that the martingale (f_n) given by $f_n = E(f|F_n)$ has $\int \sqrt{\Lambda(\lambda, f, x)}\, d\mu(x) = \infty$.*

By using the techniques in Theorem 3.7, Proposition 3.8, and Proposition 3.10, this result for martingales can be adapted to the ergodic theory case. In particular, we have this type of result.

Theorem 3.13 *Given any ergodic transformation τ, there is $f \in L_1(X)$ and some $\lambda > 0$, such that for the ergodic averages the associated jump function $\sqrt{\Lambda(\lambda, f, x)}$ has $\int \sqrt{\Lambda(\lambda, f, x)}\, d\mu(x) = \infty$.*

4 Counting the Trend To Zero

There is a very interesting counting problem that has been considered by Assani in Assani [2, 3] with some success. It is connected to some degree to the issues considered here, especially those in Section 2. Let τ be an invertible measure-preserving ergodic transformation of a probability space (X, β, μ). Let $f \in L_1(X)$, $f \geq 0$ and let

$$N_\alpha f(x) = \#\left\{k \geq 1 : \frac{f(\tau^k x)}{k} \geq \alpha\right\}.$$

The basic question is whether $\alpha N_\alpha f(x)$ tends to $\int_X f\, d\mu$ a.e. and in norm as α tends to 0? In Assani [2], this is shown to be the case in L_p, $1 < p < \infty$; the a.e. convergence was extended to the class $L \log L$ in Assani [3]. For positive functions $f \in L_1(X)$, the convergence in L_1-norm also holds. The open problem is whether a.e. convergence holds for L_1. This question reduces to the question of whether the maximal function $Cf(x) = \sup_{\alpha>0} \alpha N_\alpha f(x)$ is weak type $(1, 1)$.

Let us make one observation that may be useful in resolving this open problem, and certainly helps in working with related counting problems. For fixed $f \in L_1(X)$, consider $U_f = 1_{\{0\leq f<1\}} + \sum_{l=0}^{\infty} 2^{l+1} 1_{\{2^l\leq f<2^{l+1}\}}$ and $L_f = \sum_{l=0}^{\infty} 2^l 1_{\{2^l\leq f<2^{l+1}\}}$. Then $L_f \leq f \leq U_f$ and $2\|L_f\|_1 + \mu(\{0 \leq f < 1\}) \geq \|U_f\|_1$. So in making estimates of Cf with the intention of getting a weak type $(1, 1)$ inequality, we can work with either U_f or L_f in place of f.

Let us use this principle to arrive at several results. We want to deal also with somewhat more general versions of $N_\alpha f$. Fix a nondecreasing sequence $\nu = (\nu_k)$ and let $N_\alpha^\nu f = \#\{k \geq 1 : \frac{f(\tau^{\nu_k} x)}{k} \geq \alpha\}$. We still use the notation $N_\alpha f$ in the case that $\nu_n = n$ for all n.

Now we can estimate that

$$N_\alpha^\nu f(x) = \sum_{k=1}^{\infty} 1_{\left\{\frac{f\circ\tau^{\nu_k}}{k}\geq\alpha\right\}}(x) \leq \sum_{k=1}^{\infty} 1_{\left\{\frac{U_f\circ\tau^{\nu_k}}{k}\geq\alpha\right\}}(x)$$

because $f \le U_f$. But also, if g_1 and g_2 are disjointly supported functions in $L_1(X)$, then for any k,

$$1_{\left\{\frac{(g_1+g_2)\circ\tau^k}{k}\ge\alpha\right\}} = 1_{\left\{\frac{g_1\circ\tau^k}{k}\ge\alpha\right\}} + 1_{\left\{\frac{g_2\circ\tau^k}{k}\ge\alpha\right\}}.$$

Hence, denoting $\{2^l \le f < 2^{l+1}\}$ by E_l, for $l \ge 0$, and $E_{-1} = \{0 \le f \le 1\}$,

$$\begin{aligned}
N_\alpha^\nu f(x) &\le \sum_{k=1}^\infty \sum_{l=-1}^\infty 1_{\left\{\frac{2^{l+1}1_{E_l}\circ\tau^{\nu_k}}{k}\ge\alpha\right\}}(x) \\
&= \sum_{l=-1}^\infty \#\left\{k \ge 1 : \frac{2^{l+1}1_{E_l}(\tau^{\nu_k}x)}{k} \ge \alpha\right\} \\
&= \sum_{l=-1}^\infty \#\left\{1 \le k \le 2^{l+1}/\alpha : \frac{2^{l+1}1_{E_l}(\tau^{\nu_k}x)}{k} \ge \alpha\right\} \\
&= \sum_{l=-1}^\infty \#\{1 \le k \le 2^{l+1}/\alpha : \tau^{\nu_k}x \in E_l\}.
\end{aligned}$$

Let $A_m^\nu f(x)$ denote the usual average in ergodic theory, $A_m^\nu f(x) = \frac{1}{m}\sum_{k=1}^m f(\tau^{\nu_k}x)$, and let $A_m f(x) = A_m^\nu f(x)$ when all $\nu_n = n$. The above estimate shows that

$$N_\alpha^\nu f(x) \le \sum_{l=0}^\infty \lfloor 2^{l+1}/\alpha\rfloor A^\nu_{\lfloor 2^{l+1}/\alpha\rfloor} 1_{E_l}(x).$$

Actually, by using L_f instead of U_f, we can see that this overestimate for $N_\alpha^\nu f(x)$ is essentially an underestimate too in the sense that

$$N_\alpha^\nu f(x) \ge \sum_{l=0}^\infty \lfloor 2^{l}/\alpha\rfloor A^\nu_{\lfloor 2^{l}/\alpha\rfloor} 1_{E_l}(x).$$

Consequently, it is not hard to see that we have the following proposition.

Proposition 4.1 *The following are equivalent in the case of an ergodic transformation:*
1) For every positive $f \in L_1(X)$,

$$\lim_{\alpha\to0^+} \alpha N_\alpha f(x) = \int_X f\,d\mu \quad \text{a.e.},$$

2) For every positive $f \in L_1(X)$,

$$\sup_{\alpha>0} \sum_{l=0}^\infty 2^l A_{\lfloor 2^l/\alpha\rfloor}(1_{\{2^l\le f<2^{l+1}\}})(x) < \infty \quad \text{a.e.}$$

https://doi.org/10.4153/CJM-1999-044-2 Published online by Cambridge University Press

This result is dependent on the Pointwise Ergodic Theorem itself since $\alpha N_\alpha 1_E$ is asymptotically the averages in the Pointwise Ergodic Theorem. However, for more general sequences v, there may be no pointwise convergence result. But given some restricted information about averages along v, one can deal with $N_\alpha^v f(x)$ for suitable functions. The next result is along these lines and gives the result in Assani [2, 3] at least for L_p, $1 < p < \infty$.

Proposition 4.2 *Let $v = (v_n)$ be a sequence for which one knows the following two facts:*

1) $\lim_{m\to\infty} A_n^v 1_E$ *exists a.e. for all $E \in \beta$.*
2) for some p, $1 \le p < \infty$, and some constant C, one has for any λ and any set $E \in \beta$,

$$m\{\sup_{m\ge 1} A_m^v 1_E \ge \lambda\} \le \frac{C}{\lambda^p} m(E).$$

Then for any r, $p < r < \infty$, and any positive function $f \in L_r(X)$, $\lim_{\alpha\to 0^+} \alpha N_\alpha^v f(x)$ exists a.e.

Proof As in Assani [2, 3], it suffices to show that for any positive function $f \in L_r(X)$, we have $\sup_{\alpha>0} \alpha N_\alpha^v f(x)$ is finite a.e. We use the estimate above that gives

$$\sup_{\alpha>0} \alpha N_\alpha^v f \le C \sum_{l=-1}^{\infty} 2^l \sup_{m\ge 1} A_m^v 1_{E_l}.$$

We can see this series is finite a.e. as follows. For each l, let $c_l = \frac{1}{2^{lr/p}}$. Then by 2) above, for $f \in L_r(X)$,

$$\begin{aligned}\sum_{l=0}^{\infty} m\{\sup_{m\ge 1} A_m^v 1_{E_l} \ge c_l\} &\le C \sum_{l=0}^{\infty} 2^{rl} m(E_l) \\ &\le C\|f\|_r^r < \infty.\end{aligned}$$

Hence, for a.e. x, for large enough l, $\sup_{m\ge 1} A_m^v 1_{E_l}(x) \le c_l$. But $\sum_{l=0}^{\infty} 2^l c_l$ converges since $r > p$. Hence, $\sum_{l=-1}^{\infty} 2^l \sup_{m\ge 1} A_m^v 1_{E_l}$ converges a.e. too. ■

Remark 4.3 The question here which is analogous to Assani's problem is whether one can take $f \in L_p(X)$ and still get a.e. convergence.

A related problem suggested by the square function results in Section 2 occurs if one replaces the term $f(\tau^n x)$ in the above by a general moving averages $M_n f(x)$, with lengths L_n that are non-decreasing. Indeed, for each positive $f \in L_1(X)$, consider the operator $\Gamma_\alpha f(x) = \alpha\#\{n \ge 1 : \frac{M_n f(x)}{n} \ge \alpha\}$. Theorem 2.1 shows that $\|\Gamma_\alpha f\|_1 \le C\|f\|_1$. This fact and Assani's result in [2] suggest the question, for L_n increasing to ∞, whether $\Gamma_\alpha f(x)$ converges a.e. and in norm as α tends to 0? This question remains unanswered. But to prove the convergence in L_1-norm for this quantity, a standard approach would involve first finding an L_1-norm dense subset G in the positive functions in $L_1(X)$ for which $M_n f(x)$ converges a.e. for any $f \in G$. It is not clear if such a set G exists.

Nonetheless, the main results of Section 2 do provide some information about the level count in $\Gamma_\alpha f(x)$. Indeed, it follows directly from Corollary 2.7 that we have this rate result.

Proposition 4.4 *For any $p, 1 < p < \infty$, and for any positive function $f \in L_1(X)$,*

$$\lim_{\alpha \to 0^+} \alpha^p \# \left\{ n \geq 1 : \frac{M_n f(x)}{n} \geq \alpha \right\} = 0 \quad a.e.$$

Proof For each fixed $N \geq 1$, we have

$$\begin{aligned} \alpha^p \# \left\{ n \geq 1 : \frac{M_n f(x)}{n} \geq \alpha \right\} &\leq \alpha^p N + \alpha^p \# \left\{ n \geq N+1 : \frac{M_n f(x)}{n} \geq \alpha \right\} \\ &\leq \alpha^p N + \sum_{n=N}^{\infty} \left(\frac{M_n f(x)}{n} \right)^p. \end{aligned}$$

The proposition follows immediately from this estimate and Corollary 2.7. ■

The same principle can be coupled with Corollary 2.8 to get a rate result for a generalization of Assani's problem, where the powers of τ are allowed to be arbitrary.

Proposition 4.5 *Let (v_k) be a sequence of integers and fix $p, 1 < p < \infty$. Then for any positive function $f \in L_1(X)$,*

$$\lim_{\alpha \to 0^+} \alpha^p \# \left\{ k \geq 1 : \frac{f(\tau^{v_k} x)}{k} \geq \alpha \right\} = 0 \quad a.e.$$

Remark 4.6 If the powers (v_k) are lacunary, then even for characteristic functions $\alpha \# \{ k \geq 1 : \frac{f(\tau^{v_k} x)}{k} \geq \alpha \}$ will not converge a.e. as α tends to 0. So this result gives control of the degree of divergence for this type of level counting. It is probably best possible, but that is not completely clear yet.

There is another interesting connection between the behavior of $C|f|$ and the behavior of the p-norms of Section 2. Consider the norms $\pi_p f(x) = (\sum_{n=1}^{\infty} |\frac{f(\tau^n x)}{n}|^p)^{\frac{1}{p}}$. Let $\Pi f(x) = \sup_{1<p\leq 2} (p-1)\pi_p f(x)$. Assani in [2] and [3] considered the behavior of $(p-1)\pi_p f(x)$ as p approaches 1. He poses there the question of what happens a.e. with L_1 functions. This is not at all clear yet. In Assani [3] there is an answer to this question in case $f \in L \log L$, analogous to the results that Assani has proven about $C|f|(x)$ in this case. However, we can say that if $C|f|(x)$ satisfies a weak $(1,1)$ estimate in $L_1(X)$, then so does $\Pi f(x)$. It would also then follow that for ergodic transformations and $f \in L_1(X)$, one has $\lim_{p \to 1^+} (p-1)\pi_p f(x) = \int_X |f| \, d\mu$ in L_1-norm and a.e. too. To see this we need only observe the following result.

Proposition 4.7 *If for $f \in L_1(X)$, $C|f|(x)$ is finite a.e., then $\Pi f(x)$ is finite a.e.*

https://doi.org/10.4153/CJM-1999-044-2 Published online by Cambridge University Press

Proof Assume that f is positive. Estimate as follows:

$$
\begin{aligned}
(p-1)\sum_{n=1}^{\infty}\left|\frac{f(\tau^n x)}{n}\right|^p &= (p-1)\sum_{n=1}^{\infty}\left|\frac{f(\tau^n x)}{n}\right|^p\left(\sum_{m=0}^{\infty} 1_{\{\frac{1}{2^{m+1}}\le\frac{f\circ\tau^n}{n}<\frac{1}{2^m}\}}(x)+1_{\{1\le\frac{f\circ\tau^n}{n}\}}(x)\right)\\
&= (p-1)\sum_{n=1}^{\infty}\left|\frac{f(\tau^n x)}{n}\right|^p\left(\sum_{m=0}^{\infty} 1_{\{\frac{1}{2^{m+1}}\le\frac{f\circ\tau^n}{n}<\frac{1}{2^m}\}}(x)\right)\\
&\quad + (p-1)\sum_{n=1}^{\infty}\left|\frac{f(\tau^n x)}{n}\right|^p 1_{\{1\le\frac{f\circ\tau^n}{n}\}}(x).
\end{aligned}
$$

Let the first sum in this last expression be denoted by $\Sigma_1(p,x)$ and the second sum be denoted by $\Sigma_2(p,x)$. For a.e. x, we have $\frac{f(\tau^n x)}{n}\ge 1$ only finitely many times. Hence $\sup_{1<p\le 2}\Sigma_2(p,x)$ is finite a.e. for any $f\in L_1(X)$. But also

$$
\begin{aligned}
\Sigma_1(p,x) &\le (p-1)\sum_{m=0}^{\infty}\sum_{n=1}^{\infty}\frac{1}{2^{mp}}1_{\{\frac{1}{2^{m+1}}\le\frac{f\circ\tau^n}{n}<\frac{1}{2^m}\}}(x)\\
&\le (p-1)\sum_{m=0}^{\infty}\frac{1}{2^{mp}}\#\left\{n:\frac{f(\tau^n x)}{n}\ge\frac{1}{2^{m+1}}\right\}\\
&\le (p-1)\sum_{m=0}^{\infty}\frac{2^{m+1}}{2^{mp}}Cf(x).
\end{aligned}
$$

Since $\sum_{m=0}^{\infty}\frac{2^{m+1}}{2^{mp}}=\sum_{m=0}^{\infty}\frac{1}{2^{m(p-1)}}\le 3\frac{1}{p-1}$, this shows that $\sup_{1<p\le 2}\Sigma_1(p,x)$ is finite a.e. x whenever $Cf(x)$ is finite a.e. ■

Remark 4.8 It is not clear whether either $\Pi f(x)$ or $C|f|(x)$ is finite a.e. for every $f\in L_1(X)$. However, an inspection of the proofs of the finiteness a.e. of the norms $\pi_p f(x)$ for $f\in L_1(X)$ shows that the factor $p-1$ is certainly the correct factor to use as p approaches 1.

References

[1] M. Akcoglu, R. L. Jones and J. Rosenblatt, *The worst sums in ergodic theory*. (preprint).

[2] I. Assani, *Strong laws for weighted sums of independent identically distributed random variables*. Duke Math. J. (2) **88**(1997), 217–246.

[3] ———, *Convergence of the p-series for stationary sequences*. New York Journal of Math. **3A**(1997), 15–30, in the Proceedings of the New York Journal of Mathematics Conference, June 9–13, 1997.

[4] E. Bishop, *An upcrossing inequality with applications*. Michigan Math. J. **13**(1966), 1–13.

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

[6] J. Bourgain, *Pointwise ergodic theorems for arithmetic sets*. Inst. Hautes Études Sci. Publ. Math. **69**(1989), 5–45.

[7] D. Burkholder and R. Gundy, *Extrapolation and interpolation of quasi-linear operators on martingales*. Acta Math. **124**(1970), 249–304.

[8] J. Doob, *Stochastic Processes*. John Wiley & Sons, Inc., New York, Chapman & Hall, Limited, London, 1953.

[9] A. Garsia, *Topics in Almost Everywhere Convergence*. Lectures in Advanced Mathematics, Markham Publishing, Chicago, 1970.

[10] V. V. Ivanov, *Geometric properties of monotone functions and probabilities of random fluctuations*. Siberian Math. J. (1) **37**(1996), 102–129.
[11] R. L. Jones, *Inequalities for the ergodic maximal function*. Studia Math. **40**(1977), 111–129.
[12] R. L. Jones, I. Ostrovskii and J. Rosenblatt, *Square functions in ergodic theory*. Ergodic Theory Dynamical Systems **16**(1996), 267–305.
[13] R. L. Jones, R. Kaufman, J. Rosenblatt, and M. Wierdl, *Oscillation in ergodic theory*. Ergodic Theory Dynamical Systems **18**(1998), 889–935.
[14] A. G. Kachurovskii, *The rate of convergence in ergodic theorems*. Russian Math. Surveys (4) **51**(1996), 653–703.
[15] S. Kalikow and B. Weiss, *Fluctuations of ergodic averages*. (preprint).
[16] J. Rosenblatt and M. Wierdl, *Fourier Analysis and Almost Everywhere Convergence*. Proceedings of the Conference on Ergodic Theory and its Connections with Harmonic Analysis, Alexandria, Egypt, Cambridge University Press, 1994, 3–151.

Department of Mathematics
DePaul University
2219 N. Kenmore
Chicago, Illinois 60614
U.S.A.
e-mail: rjones@condor.depaul.edu

Department of Mathematics
University of Illinois at Urbana
Urbana, Illinois 61801
U.S.A.
e-mail: jrsnbltt@math.uiuc.edu

Department of Mathematical Sciences
University of Memphis
Memphis, Tennessee 38152
U.S.A.
e-mail: mw@wierdlmpc.msci.memphis.edu