JOURNAL OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 00, Number 0, Pages 000–000
S 0894-0347(XX)0000-0

RATES OF DIVERGENCE OF NONCONVENTIONAL ERGODIC AVERAGES

ANTHONY QUAS AND MÁTÉ WIERDL

Contents

1. Introduction

As the statements of many of the results below are general but fairly technical, we start by stating some concrete results that are formally corollaries of the main theorems, but which should give the reader a good idea about our subject. Our choice of discussing these special results is all the more justified, since they, in fact, motivated the work in the paper.

We consider ergodic sums along a sequence (a_n): $S_N f(x) = \sum_{n=1}^{N} f(T^{a_n}x)$ and ask for the maximal growth rate of these sums.

Theorem A.

(1) *Let $f \in L^1$, let T be a measure-preserving transformation and let (a_n) be an arbitrary sequence. Then*

$$\lim_{N\to\infty} \frac{1}{N \log N \log\log N \ldots} \sum_{n=1}^{N} f(T^{a_n}x) = 0 \text{ a.e.},$$

where the product in the denominator is taken over those terms that exceed 1.

(2) *Let T be an aperiodic measure-preserving transformation and let the sequence (M_N) satisfy $M_N/(N \log N \log\log N \ldots) \to 0$. Then there exists an*

2000 *Mathematics Subject Classification.* 37A30.

This research was supported by research grants from the NSF.

$f \in L^1$ and a sequence (a_n) such that

$$\limsup_{N\to\infty} \frac{1}{M_N} \sum_{n=1}^{N} f(T^{a_n}x) = \infty \text{ a.e..}$$

Further, the sequence (a_n) may be taken to be the sequence (2^n).

Remark. In [2], Akcoglu, Jones, and Rosenblatt proved that if $\sum_{N=1}^{\infty} 1/M_N$ is finite, then $(1/M_N)S_N f$ is convergent for $f \in L^1$ and also it was demonstrated that if M_N is taken to be any sequence of the form $N \log N \ldots \log^{(k)} N$ (where $\log^{(k)}$ denotes the k-fold composition of log), then there exists $f \in L^1$ for which $(1/M_N)S_N f$ is divergent. Based on this, they conjectured that $(1/M_N)S_N f$ is convergent if and only if $\sum_N 1/M_N$ is finite. However, one can check that the example $M_N = N \log N \log\log N \ldots$ disproves this conjecture.

Theorem B. *Let $p > 1$.*

(1) *Let $f \in L^p$, let T be a measure-preserving transformation and let (a_n) be an arbitrary sequence. Then*

$$\lim_{N\to\infty} \frac{1}{N(\log N)^{1/p}} \sum_{n=1}^{N} f(T^{a_n}x) = 0 \text{ a.e..}$$

(2) *Let T be an aperiodic measure-preserving transformation and let the sequence (M_N) satisfy $M_N/(N(\log N)^{1/p}) \to 0$. Then there exists an $f \in L^p$ and a sequence (a_n) such that*

$$\limsup_{N\to\infty} \frac{1}{M_N} \sum_{n=1}^{N} f(T^{a_n}x) = \infty \text{ a.e..}$$

Further, the sequence (a_n) may be taken to be the sequence (2^n).

For another application of the general techniques presented in the first part of the paper, we study some averages introduced by Khintchine: For $f \in L^1([0,1))$, we let $K_N f(x) = \sum_{n \le N} f(nx \bmod 1)$. Khintchine [8] in 1923 conjectured that $(1/N)K_N f(x)$ converges for almost every x to the integral of f. This was shown to be false by Marstrand [12] in 1970. Later, Bourgain [4] gave an alternative proof using his entropy method. In Section 5, we give a very simple and brief demonstration of Marstrand's result using Rokhlin towers. In fact, we show more: we demonstrate that for suitable $f \in L^p$, the growth rate of $K_N f$ is exactly the same as the maximal growth rate obtained in Theorem B.

These techniques also allow us to resolve a question of Nair [13] concerning the Khinchine averages taken along a multiplicative subsemigroup of the natural numbers rather than all natural numbers. Our results demonstrate that the averages $(1/|G \cap [1,N]|) \sum_{\{n \in G:\, n \le N\}} f(nx \bmod 1)$ converges for $f \in L^1$ to the integral if and only if the semigroup G is a subsemigroup of one that is finitely generated.

We would like to thank Ciprian Demeter for making available to us his preprint [5]. Many of the ideas in that paper were crucial to us in formulating the results of Section 4. We would also like to thank Michael Boshernitzan, Roger Jones and Joe Rosenblatt for stimulating discussions.

2. Background and Statement of Results

We will make extensive use in what follows of so-called weak ($L^{p,\infty}$) norms. Given a function f on a measure space (X, μ), its $L^{p,\infty}$ norm is defined by

$$\|f\|_{p,\infty} = \sup_y y \cdot \mu\{x\colon |f(x)| \geq y\}^{1/p}.$$

In the case of a sequence (w_t), its norm analagously is $\|w\|_{p,\infty} = \sup_y y \cdot |\{t\colon w_t \geq y\}|^{1/p}$. As is well-known, these "norms" fail to be sub-additive. In the case $p > 1$, there is a true norm $|||\cdot|||_{p,\infty}$ and a constant $C > 1$ such that $\|\cdot\|_{p,\infty}/C \leq |||\cdot|||_{p,\infty} \leq C\|\cdot\|_{p,\infty}$. In the case $p = 1$ however, there is no equivalent norm. For more details about these norms, the reader is referred to Bennett and Sharpley's book [3].

In this paper, we will consider almost everywhere convergence of sequences of the form $w_t A_t f(x)$, where (w_t) is a sequence of real numbers and the A_t are averaging operators of various kinds. The typical example that we will consider is the case $A_t f(x) = (1/2^t) \sum_{n \leq 2^t} f(T^{a_n} x)$, where (a_n) is a sequence of times and T is a measure-preserving transformation. A key tool in our work will be the maximal operator $(wA)^* f(x) = \sup_t w_t A_t f(x)$. We will say that the sequence of operators $(w_t A_t)$ *satisfies a weak (p,p) maximal inequality* if there exists a constant $C > 0$ such that $\|(wA)^* f\|_{p,\infty} \leq C\|f\|_p$ for all $f \in L^p$.

Fact 2.1. *Under conditions that are satisfied by all of the operators that we consider in this paper, we have the following:*

(1) *(Banach Principle) If the sequence $(w_t A_t)$ satisfies a weak (p,p) maximal inequality, then the set of functions $f \in L^p$ for which $w_t A_t f(x)$ is convergent almost everywhere is a closed set in L^p.*
(2) *If the sequence $(w_t A_t)$ fails to satisfy a weak (p,p) maximal inequality in one measure-preserving system, then there is a function $f \in L^p$ such that $\limsup_{t\to\infty} w_t A_t f(x) = \infty$ almost everywhere in any of the systems that we consider.*

The first statement is well known (see for example Rosenblatt and Wierdl's article [15] or Garsia's book [6]), holding under very mild conditions on the operator. Since in this paper, convergence will hold trivially on the dense set of bounded measurable functions, in order to prove a positive result, it will be sufficient to establish a maximal inequality

The second statement is based on the transference principle of Calderón, a theorem of Sawyer [16] and an adaptation appearing in an article of Akcoglu, Bellow, Jones, Losert, Reinhold-Larsson and Wierdl [1]. First, the transference principle tells us that if a maximal inequality fails in one measure-preserving system along some given sequence of times, then the maximal inequality fails in all measure-preserving systems along the sequence of times. Sawyer proves that if a sequence of operators on a finite measure space fails to satisfy a maximal inequality and commutes with a "mixing family" of transformations, then there exists a function giving divergence almost everywhere. The paper [1] reaches the same conclusion for a family of operators that are averages of iterates of a single aperiodic measure-preserving transformation.

We will use Iverson notation for indicator functions so that by the expression $[y < w_t < 2^t y]$, we will mean the function that is equal to 1 when the condition is satisfied and 0 otherwise. For a sequence $(w_t)_{t\in\mathbb{N}}$ of positive real numbers, define

$$C_p(w) = \begin{cases} \sup_y \sum_t [y < w_t < 2^t y] w_t & \text{for } p = 1 \\ \|w\|_{p,\infty} & \text{for } 1 < p < \infty. \end{cases}$$

We note that it is convenient to formulate the results not in terms of the ergodic sums up to N as was done in the introduction, but rather to consider the ergodic sums (or equivalently ergodic averages) up to 2^t. We justify this restriction as follows. First, we observe that it is sufficient to establish convergence to 0 for non-negative functions. Let (u_n) be a sequence of real numbers and let $v_t = \max_{2^{t-1}<n\le 2^t} u_n$. We will show that the following three statements are equivalent:

(1) $u_N \sum_{n\le N} f(T^{a_n}x) \to 0$ a.e. x, for all $f \in L^p$ and every sequence (a_n).
(2) $v_t \sum_{n\le 2^{t-1}} f(T^{a_n}x) \to 0$ a.e. x, for all $f \in L^p$ and every sequence (a_n).
(3) $v_t \sum_{n\le 2^t} f(T^{a_n}x) \to 0$ a.e. x, for all $f \in L^p$ and every sequence (a_n).

To see this, note that for any t, x and N satisfying $2^{t-1} < N \le 2^t$,

$$v_t \sum_{n\le 2^{t-1}} f(T^{a_n}x) \le 2u_N \sum_{n\le N} f(T^{a_n}x) \le 4v_t \sum_{n\le 2^t} f(T^{a_n}x).$$

It follows that (3) implies (1) implies (2).

Suppose finally that (2) is satisfied. Let (v_t), (a_n) and $f \in L^p$ be given. Let $b_n = a_{2n-1}$ and $b'_n = a_{2n}$. Applying (2) separately to the sequences (b_n) and (b'_n) and summing, we deduce (3).

Theorem 2.2. *Let (w_t) be a sequence of positive real numbers such that $C_1(w) < \infty$. For each $t \in \mathbb{N}$, let the set $\mathcal{T}_t$ contain at most 2^t measure-preserving transformations. Then for $f \in L^1$,*

$$\lim_{t\to\infty} w_t 2^{-t} \sum_{T\in\mathcal{T}_t} f(Tx) = 0$$

almost everywhere.

Theorem 2.3. *Let $1 < r < p < \infty$ and let (w_t) be a sequence of positive real numbers such that $\|w\|_{p,\infty} < \infty$. For each $t \in \mathbb{N}$, let A_t be an $L^r - L^\infty$ contraction. Then for any $f \in L^p$,*

$$\lim_{t\to\infty} w_t A_t f(x) = 0$$

almost everywhere.

In particular, if for each $t \in \mathbb{N}$, the set $\mathcal{T}_t$ contains at most 2^t measure-preserving transformations. Then for $f \in L^p$,

$$\lim_{t\to\infty} w_t 2^{-t} \sum_{T\in\mathcal{T}_t} f(Tx) = 0$$

almost everywhere.

Theorem 2.4. *Let $1 \le p < \infty$ and let (w_t) be a sequence of positive real numbers such that $C_p(w) = \infty$. Let T be an aperiodic probability measure-preserving transformation. Then there is a sequence (a_n) of integers so that the maximal function of the averages*

$$w_t 2^{-t} \sum_{n\le 2^t} f(T^{a_n}x)$$

is not weak (p,p) *and hence there exists an* $f \in L^p$ *for which the averages diverge almost everywhere.*

The following proposition gives a simple description of the w_t for which $C_p(w_t) < \infty$ in the case that the w_t are a sufficiently regularly decaying sequence.

Proposition 2.5. *Let* (w_t) *be a sequence of weights and let* $\Phi(t) = t \log t \log\log t \dots$ *be defined to be the product of* t *and all iterates of* log *that are defined and greater than 1 at* t*. Let* $1 < p < \infty$*.*

1. *If there exists a* K *such that* $w(t) \le K/\Phi(t)$*, then* $C_1(w) < \infty$*.*
2. *If* $w(t)\Phi(t) \to \infty$ *as* $t \to \infty$*, then* $C_1(w) = \infty$*.*
3. *If there exists a* K *such that* $w(t) \le Kt^{-1/p}$*, then* $C_p(w) < \infty$*.*
4. *If* $w(t)t^{1/p} \to \infty$ *as* $t \to \infty$*, then* $C_p(w) = \infty$*.*

Remark 2.6.

- Notice that Theorems 2.2, 2.3 and 2.4 give a dichotomy: in any L^p, $(p \ge 1)$, if $C_p(w)$ is finite then the averages $w_t 2^{-t} \sum_{n<2^t} f(T^{a_n}x)$ converge along all sequences of times (a_n) for all L^p functions f, whereas if $C_p(w)$ is infinite then in every aperiodic dynamical system there exists a sequence of times (a_n) and an L^p function f for which the averages fail to converge.
- We strengthen this dichotomy below by showing that there are sequences of times (a_n) that can be chosen independently of w such that if $C_p(w) = \infty$, then in every aperiodic dynamical system, there exists an f in L^p such that $w_t 2^{-t} \sum_{n<2^t} f(T^{a_n}x)$ diverges almost everywhere.
- Theorem 2.2 fails in the L^1 case if the transformations are taken to be $L^1 - L^\infty$ contractions. Also Theorem 2.3 fails in the L^p case if the contractions are only assumed to be $L^p - L^\infty$ contractions.

3. Proofs of maximal rate theorems

We make the following observations concerning the relationships of L^p goodness for various L^p:

If $1 < p, q < \infty$, then $C_p(w) < \infty$ if and only if $C_q(w^{p/q}) < \infty$. To see this, note that $C_p(w) < \infty$ if and only if $\|w^p\|_{1,\infty} < \infty$ if and only if $\|w^{p/q}\|_{q,\infty} < \infty$.

If $C_1(w) < \infty$ then $C_p(w^{1/p}) < \infty$ for $p > 1$. To see this, we note that $w_n \le C_1(w)$ for all n and argue as follows:

$$\begin{aligned} y\#\{n\colon w_n > y\} &= y\#\{n\colon w_n \ge 2^n y\} + y\#\{n\colon y < w_n < 2^n y\} \\ &\le y\#\{n\colon 2^n \le C_1(w)/y\} + \sum_n [y < w_n < 2^n y] w_n \\ &\le 2C_1(w). \end{aligned}$$

This shows that $\|w\|_{1,\infty} \le 2C_1(w)$ so that $C_p(w^{1/p}) = \|w^{1/p}\|_{p,\infty} = \|w\|_{1,\infty}^{1/p} < \infty$. The converse to this assertion fails as is seen by considering $w_t = 1/t$.

Proof of Proposition 2.5. We deal first with the equivalence $\Phi(t)w_t$ is bounded above if and only if $C_1(w) < \infty$.

We start by defining a quantity $C_1'(w)$ such that $C_1(w) = \infty$ if and only if $C_1'(w) = \infty$. Namely, define

$$C_1'(w) = \sup_z \sum_t [t > z \text{ and } w_t > 2^{-z}] w_t.$$

Writing $y = 2^{-z}$, this may be rewritten $C_1'(w) = \sup_y \sum_t [y < w_t \text{ and } 1 < 2^t y] w_t$. Comparing with $C_1(w) = \sup_y \sum_t [y < w_t < 2^t y] w_t$, we see that

$$|C_1'(w) - C_1(w)| \le \sum_t [w_t < 2^t y \le 1] w_t + \sum_t [w_t \ge 2^t y > 1] w_t.$$

If $\limsup w_t > 0$, it is easy to see that both $C_1(w)$ and $C_1'(w)$ are infinite. Otherwise, since there are only finitely many terms with $w_t > 1$, the second term in the above inequality is finite. The first term is bounded above by $\sum_t [2^t y \le 1] 2^t y \le 2$ showing that $|C_1'(w) - C_1(w)| < \infty$ as required.

Since $\Phi(t)/t \to \infty$ and $\Phi(t/(\log t)^2)/t \to 0$, we see that for large y, $2^y/y^2 < \Phi^{-1}(2^y) < 2^y$.

We consider $\sum_{y<t<\Phi^{-1}(2^y)} 1/\Phi(t)$. A calculation by comparison with the integral shows that

$$\begin{aligned} \sum_{y<t<\Phi^{-1}(2^y)} \frac{1}{\Phi(t)} &\sim \sum_{y<t<\Phi^{-1}(2^y)} \frac{1}{t \log t \log\log t \ldots} \\ &\sim \int_y^{\Phi^{-1}(2^y)} \frac{1}{t \log t \log\log t \ldots} \, dt \\ &\sim 1. \end{aligned} \tag{1}$$

If $\Phi(t) w_t \to \infty$, we see that for large t, $1/\Phi(t) > 2^{-y}$ implies $w_t > 2^{-y}$ so

$$\begin{aligned} \sum_{\{t\colon t>y \text{ and } w_t > 2^{-y}\}} w_t &\ge \sum_{\{t\colon t>y \text{ and } \Phi(t)<2^y\}} w_t \\ &= \sum_{y<t<\Phi^{-1}(2^y)} \frac{1}{\Phi(t)} w_t \Phi(t). \end{aligned}$$

From equation (1), we see that this is divergent establishing part (2) of the proposition.

If on the other hand, $w_t \Phi(t)$ is bounded above, we have $w_t \le k/\Phi(t)$. The above calculation then shows that $C_1'(w) < \infty$ establishing part (1).

For the L^p case, Suppose that $t^{1/p} w_t \to \infty$. Given any M, for t greater than some t_0, $w_t > M/t^{1/p}$. It follows that for sufficiently small s, the number of t such that $w_t > s$ is at least M^p/s^p. Since M is arbitrary, it follows that $C_p(w) = \infty$. This proves part (4). If on the other hand, $t^{1/p} w_t$ is bounded above, we have $w_t < k/t^{1/p}$ for some k so that the number of solutions to $w_t > s$ is bounded above by k^p/s^p for all y so that $C_p(w) < k$ giving condition (3). □

In order to prove Theorem 2.2, we start with a lemma.

Lemma 3.1. *Let (w_t) be a sequence of positive numbers, let $\mathcal{T}_t$ be a set of at most 2^t measure-preserving transformations of a probability space X and denote by $A_t f(x)$, the quantity $1/2^t \sum_{T \in \mathcal{T}_t} f(Tx)$. Then for any $f \in L^1$, we have*

$$\| \sup_t w_t A_t f \|_{1,\infty} \le 9 C_1(w) \|f\|_1$$

Proof. We want to show that for every $\lambda > 0$ and $f \in L^1$,

$$\mu \left(\sup_t w_t A_t f(x) > \lambda \right) \le 9 C_1(w) \lambda \|f\|_1.$$

We will in fact prove the following apparently stronger inequality.

$$\sum_t \mu\left(w_t A_t f(x) > \lambda\right) \leq \frac{9C_1(w)}{\lambda}\|f\|_1.$$

Fix an $f \in L^1$. By rescaling f if necessary, we can assume that $\lambda = 3$. For a fixed t, let us split f into three parts, up, middle, and down: $f = u + m + d$ where

$$\begin{aligned} u &= u_t = [f \geq 2^t/w_t]f \\ m &= m_t = [1/w_t < f < 2^t/w_t]f \\ d &= d_t = [f \leq 1/w_t]f \end{aligned}$$

We first estimate the upper part, $u = u_t$. We note that the set of x where $w_t A_t u(x) > 1$ is a subset of the set of x where $A_t u(x) > 0$. The set on which $A_t u(x) > 0$ is of measure at most 2^t times the measure of the set on which u is supported. It follows that $\mu\{x\colon w_t A_t u(x) > 1\} \leq 2^t \mu\{x\colon f(x) \geq 2^t/w_t\}$.

We can check that $C_1(w) < \infty$ implies that the w_t are bounded above by $C_1(w)$. Hence summing over t, we get

$$\begin{aligned} \sum_t \mu\left(w_t A_t u_t(x) > 1\right) &\leq \sum_t 2^t \mu(f \geq 2^t/w_t) \\ &\leq \sum_t 2^t \mu(C_1(w) f \geq 2^t) \leq 2C_1(w)\|f\|_1. \end{aligned}$$

Now for the middle part, m_t,

$$\begin{aligned} \mu\left(w_t A_t m_t(x) > 1\right) &\leq \int w_t A_t m_t = w_t \int m_t \\ &= w_t \int [1/w_t < f < 2^t/w_t] f \end{aligned}$$

Summing over t, and interchanging the summation and integration

$$\begin{aligned} \sum_t \mu\left(w_t A_t m_t(x) > 1\right) &\leq \sum_t w_t \int [1/w_t < f < 2^t/w_t] f \\ &= \int f(x) \sum_t [1/f(x) < w_t < 2^t/f(x)] w_t \, d\mu(x). \end{aligned}$$

Using the assumption that $C_1(w) < \infty$ (with the y taken to be $1/f(x)$), we get

$$\sum_t \mu\left(w_t A_t m_t(x) > 1\right) \leq \int C_1(w) f = C_1(w)\|f\|_1.$$

As for the down part, clearly, for every t,

$$\left\{ w_t \frac{1}{2^t} \sum_{T \in \mathcal{T}_t} d_t(Tx) > 1 \right\} = \varnothing,$$

since $d_t \leq 1/w_t$.

Summing, we see that $\mu(\sup_t w_t A_t f > 3) < 3C_1(w)\|f\|_1$ as required. □

Proof of Theorem 2.2. The above lemma establishes a maximal inequality. Since there is a dense class of bounded functions on which there is almost everywhere convergence, it follows that there is convergence for all $f \in L^1$ as required. □

Before proving Theorem 2.3, we prove a general lemma that will imply the theorem almost immediately.

Lemma 3.2. *Let $1 \le r < p < \infty$, let (w_t) be a sequence of positive real numbers and A_t be a sequence of positive L^r-L^∞ contractions. Then there is a constant C depending on p and r such that*

$$\left\|\sup_t w_t A_t f\right\|_{p,\infty} \le C\|f\|_p\|w\|_{p,\infty}.$$

Proof. We need to estimate $\sup_\lambda \lambda^p \mu\{\sup_t w_t A_t f > \lambda\}$. Since the inequality is homogeneous in f, it is sufficient to prove the estimate in the case that $\lambda = 2$. For a fixed n, we write f as the sum of u_t and d_t, where $d_t = [f \le 1/w_t]f$ and $u_t = [f > 1/w_t]f$. Since A_t is an L^∞ contraction, we see that $w_t A_t d_t \le 1$ so that a necessary condition for $w_t A_t f > 2$ is $w_t A_t u_t > 1$. We then estimate as follows:

$$\begin{aligned}
&\mu\{\sup_t w_t A_t f > 2\} \le \mu\{\sup_t w_t A_t u_t > 1\} \\
&\le \sum_t \mu\{A_t u_t > 1/w_t\} = \sum_t \mu\{(A_t u_t)^r > w_t^{-r}\} \\
&\le \sum_t w_t^r \int (A_t u_t)^r \, d\mu \le \sum_t w_t^r \int u_t^r \, d\mu \\
&= \int f^r(x) \sum_t w_t^r [f(x) > 1/w_t] \, d\mu(x) \\
&= \int f^r(x) \sum_t w_t^r [w_t > 1/f(x)] \, d\mu(x)
\end{aligned}$$

Splitting the summation according into parts on which the w_t lie between consecutive powers of e^{-1}, this is further bounded above by

$$\begin{aligned}
&e^r \int f^r(x) \sum_{0 \le j < \log f(x)} \sum_{\{t\colon e^{-j} < w_t \le e^{-j+1}\}} e^{-rj} \, d\mu(x) \\
&\le e^r \int f^r(x) \sum_{0 \le j < \log f(x)} e^{-rj} \#\{t\colon w_t > e^{-j}\} \, d\mu(x) \\
&\le e^r \int f^r(x) \sum_{0 \le j < \log f(x)} e^{-rj} \|w\|_{p,\infty}^p e^{pj} \, d\mu(x) \\
&= e^r \|w\|_{p,\infty}^p \int f^r(x) \sum_{0 \le j < \log f(x)} e^{(p-r)j} \, d\mu(x) \\
&\le \frac{e^p}{e^{p-r} - 1} \|w\|_{p,\infty}^p \int f^r(x)(f(x))^{p-r} \, d\mu(x) = \frac{e^p}{e^{p-r} - 1} \|w\|_{p,\infty}^p \|f\|_p^p.
\end{aligned}$$

□

Proof of Theorem 2.3. The above lemma establishes an L^p maximal inequality. Since for bounded functions there is convergence to 0 and these form a dense subset of L^p, the theorem follows. □

4. Ultimate Badness

Definition. The sequence (T_n) of linear operators on L^p is called *ultimately bad in* L^p if for any (w_t) satisfying $C_p(w) = \infty$, the maximal function of the averages

$$w_t \frac{1}{2^t} \sum_{n \le 2^t} T_n f(x)$$

is not weak (p, p).

A sequence (a_n) of real numbers (integers) is called ultimately bad in L^p if for any aperiodic measure-preserving flow $(T^t)_{t\in\mathbb{R}}$ (aperiodic measure-preserving transformation T) the operators T^{a_n} are ultimately bad in L^p.

Remark 4.1. By Fact 2.1, in all sequences (T_n) considered in this paper for which the above averages fail to satisfy a weak inequality, there is an $f \in L^p$ such that $\limsup_{t\to\infty}(w_t/2^t)\sum_{n\le 2^t} T_n f(x)$ is infinite almost everywhere.

Remark 4.2. If a sequence of transformations has a subsequence with bounded gaps that is ultimately bad in L^p, then the original sequence is also ultimately bad in L^p.

We start the section by giving some equivalent formulations of ultimate badness of sequences of times.

Theorem 4.3. *Let $1 \le p < \infty$. The following are equivalent.*

1. *The sequence of times (a_n) is ultimately bad for L^p.*
2. *There exists a B such that for any sequence (w_t) with $C_p(w) < \infty$, there is an $f \in L^p$ such that*

$$\left\| \sup_t w_t A_t f \right\|_{p,\infty} \ge B C_p(w) \|f\|_p,$$

where $A_t f(x) = 1/2^t \sum_{j \le 2^t} f(T^{a_j} x)$.

Proof. Suppose we are given that condition (2) holds. Supposing further that $C_p(w) = \infty$, we can take truncations $w^{(n)}$ of w with $C_p(w^{(n)})$ increasing to infinity. Then letting $f^{(n)}$ be the function guaranteed by the condition, we see that $\|\sup_t w_t A_t f^{(n)}\|_{p,\infty} \ge \|\sup_t w_t^{(n)} A_t f^{(n)}\|_{p,\infty} > BC_p(w^{(n)})\|f^{(n)}\|_p$. Since the constants $BC_p(w^{(n)})$ increase to ∞, the ultimate badness follows so that condition (2) implies condition (1).

To show that condition (1) implies condition (2), we argue by the contrapositive. Suppose that no constant B as in condition (2) exists. Then for each $k \in \mathbb{N}$, there exists a sequence $(w_t^{(k)})$ such that

$$\sup_{\|f\|=1} \left\| \sup_t w_t^{(k)} A_t f \right\|_{p,\infty} \le 4^{-k} C_p(w^{(k)})$$

We may assume that the sequences $(w_t^{(k)})$ are scaled so that $C_p(w^{(k)}) = 2^k$ and $\sup_{\|f\|=1} \left\| \sup_t w_t^{(k)} A_t f \right\|_{p,\infty} \le 2^{-k}$. Forming a new sequence $v_t = \sum_k w_t^{(k)}$, we observe that

$$\sup_{\|f\|=1}\left\|\sup_t v_t A_t f\right\|_{p,\infty} = \sup_{\|f\|=1}\left\|\sup_t \sum_k w_t^{(k)} A_t f\right\|_{p,\infty} \le \sup_{\|f\|=1}\left\|\sum_k \sup_t w_t^{(k)} A_t f\right\|_{p,\infty}.$$

Since for $\|f\|_p = 1$, $\|\sup_t w_t^{(k)} A_t f\|_{p,\infty} \le 2^{-k}$, the norm of the sum is bounded above by a constant depending only on p. (In the case $p = 1$, this follows from a result of Stein and Weiss [17]). On the other hand, since $C_p(v) = \infty$, this establishes that the sequence (a_n) is not ultimately bad in L^p so that condition (1) implies condition (2). □

Theorem 4.4. *Let $1 < p < \infty$ and let (a_n) be a sequence of times. The following conditions are equivalent*

(1) *The sequence of times (a_n) is ultimately bad for L^p;*
(2) *There exists a $C > 0$ such that for all sequences of weights (w_t) such that $\|w\|_{p,\infty} < \infty$, there exists an $f \in L^p$ such that $\|\sup_t w_t A_t f\|_{p,\infty} \ge C\|w\|_{p,\infty}\|f\|_p$.*
(3) *There exists a $C > 0$ such that for all finite subsets $J \subset N$, there exists an $f \in L^p$ such that $\|\max_{j\in J} A_j f\|_{p,\infty} \ge C|J|^{1/p}\|f\|_p$.*
(4) *There exists a $C > 0$ such that for all finite subsets $J \subset N$, there exists an $f \in L^p$ such that $\|\max_{j\in J} A_j f\|_p \ge C|J|^{1/p}\|f\|_p$.*
(5) *The sequence of times (a_n) is ultimately bad for L^q for all $1 < q < \infty$.*

The equivalence of (1) and (2) was already established in Theorem 4.3. The structure of the proof is that we first prove (2) is equivalent to (3). Since $\|f\|_p \ge \|f\|_{p,\infty}$, we see that (3) implies (4). Most of the work is taken up with proving the implication (4) implies (3). The implication (4) implies (5) falls out of the proof of this step.

Remark 4.5. We remark that at no point in the proof do we use the fact that the T^{a_n} are powers of a single measure-preserving transformation. The whole proof works verbatim if the T^{a_n} are replaced by a family of measure-preserving transformations. We make use of the theorem in this form in Section 5.

Remark. We note that a further by-product of the proof is the fact that in fact if (a_n) is ultimately bad for L^p, then there is a C such that for every $J \subset \mathbb{N}$, there exists a characteristic function f such that $\|\max_{j\in J} A_j f\|_{p,\infty} \ge C|J|^{1/p}\|f\|_p$.

Remark. In the L^p case above, if we restrict to *decreasing* sequences (w_t), then condition (3) can be weakened to

$$\left\|\max_{1\le j\le N} A_j f\right\|_{p,\infty} > c\cdot N^{1/p}\cdot \|f\|_p\,.$$

This condition is the same as one appearing in recent work of Demeter [5]

Proof of Theorem 4.4: (2) *is equivalent to* (3). To see that condition (2) implies condition (3), let J be a finite subset of the positive integers and let (w_t) be the indicator sequence of the set J. It is not hard to see that $\|w\|_{p,\infty} = |J|^{1/p}$ and condition (3) follows.

Now suppose that condition (3) holds with a constant C. Let the sequence (w_t) satisfy $\|w\|_{p,\infty} < \infty$. Let a positive number $\sigma < 1$ be given and let λ be such that

$$\lambda \cdot |\{t\colon w_t > \lambda\}|^{1/p} > \sigma \cdot \|w_j\|_{p,\infty}\,.$$

Setting $J = \{j\colon w_j > \lambda\}$ the above can be written as

$$\lambda \cdot |J|^{1/p} > \sigma \cdot \|w\|_{p,\infty}\,.$$

By condition (3), there exists an f such that $\|\max_{j\in J} A_j f\|_{p,\infty} \geq C|J|^{1/p}\|f\|$. Now estimate as

$$\begin{aligned}\left\|\sup_j w_j \cdot A_j f\right\|_{p,\infty} &\geq \left\|\max_{j\in J} w_j \cdot A_j f\right\|_{p,\infty}\\ &> \left\|\max_{j\in J} \lambda \cdot A_j f\right\|_{p,\infty}\\ &\geq C\cdot\lambda\cdot|J|^{1/p}\,\|f\|_p\\ &> C\cdot\sigma\cdot\|w_j\|_{p,\infty}\,\|f\|_p\,,\end{aligned}$$

where the third inequality comes from condition (3). This shows that condition (2) follows. □

The proof of (4) implies (3) proceeds by three lemmas. An outline of the argument is as follows. Let $\|\max_{j\in J} A_j f\|_p \geq C|J|^{1/p}\|f\|_p$. We split f into pieces according to the value of $f(x)$: $f_n = f\mathbf{1}_{\{x\colon f(x)\in(R^{n-1/2},R^{n+1/2}]\}}$, where R is a suitably chosen constant. The pieces are then essentially disjointly supported multiples of characteristic functions. The fact that $\|\sup_j A_j f\|_p$ is comparable to $|J|^{1/p}\|f\|_p$ (which is the maximum possible value over all possible operators of this type) is shown to imply that the $\sup_j A_j f_n$ are almost disjointly supported; they are themselves close to characteristic functions; and many have a norm that is comparable to the maximum norm. Since for multiples of characteristic functions, the weak and strong norms are equal, this will allow us to conclude condition (3).

Lemma 4.6. *Suppose that for some finite subset J of $\mathbb{N}$ and some function $f \in L^p$, $\|\max_{j\in J} A_j f\|_p = C|J|^{1/p}\|f\|_p$. Then there exists a subset J' of J such that for $j \in J'$, $\|A_j f\| \geq (C/2)\|f\|$ and $\|\max_{j\in J'} A_j f\|_p \geq (C/2)|J|^{1/p}\|f\|_p$.*

Proof. Let $J_1 = \{j\colon \|A_j f\| < (C/2)\|f\|_p\}$ and $J' = J\setminus J_1$. For $j \in J$, let $E_j = \{x\colon A_j f(x) = \max_{k\in J} A_k f(x)\}$. We have

$$\begin{aligned}\sum_{j\in J'}\int_{E_j}(A_j f)^p &= \int \max_{j\in J}(A_j f(x))^p - \sum_{j\in J_1}\int_{E_j}(A_j f)^p\\ &\geq C^p|J|\|f\|^p - \sum_{j\in J_1}\int (A_j f)^p\\ &\geq C^p|J|\|f\|^p - |J|(C/2)^p\|f\|^p \geq (C/2)^p|J|\|f\|^p.\end{aligned} \tag{2}$$

The conclusion then follows: $\|\max_{j\in J'} A_j f\|^p \geq \sum_{j\in J'}\int_{E_j}(A_j f)^p \geq (C/2)^p|J|\|f\|^p$ □

By an *averaging operator*, A, we will mean an operator of the form $Af(x) = \frac{1}{N}\sum_{n\leq N} f(T^{a_n}x)$. Given an averaging operator A, a fixed non-negative function

f and a real number $R > 1$, let E^k denote $\{x\colon Af(x) \in (R^{k-1/2}, R^{k+1/2}]\}$. Given all of this, we define for $k \in \mathbb{Z}$,

$$B^k g(x) = \frac{1}{N} \sum_{\{n \le N\colon f(T^{a_n}x) \in (R^{k-1}, R^{k+1}]\}} g(T^{a_n}x) \mathbf{1}_{E^k}.$$

Note that the range of the summation does not depend on g so that B^k is a linear operator. We modify this giving $B'^k g(x) = B^k g(x) \mathbf{1}_{\{B^k g(x) > R^{k-1}\}}(x)$. Also define $Bg(x) = \sum_k B^k g(x)$ and $B'g(x) = \sum_k B'^k g(x)$.

Lemma 4.7. *Let $0 < C < 1$ be given. There exists an R with the following property: If A is an averaging operator such that $\|Af\|_p \ge (C/2)\|f\|$, then if B' is defined as above, we have $\|Af - B'f\| \le \|Af\|/2$.*

Proof. Let L be chosen so that $(2/C)L^{-1/p} = 1/8$ and let the quantity R in the statement of the lemma be chosen so that $\max(L/R^{(p-1)/2}, R^{-1/2}) = 1/8$. Note that R depends only on C.

First define

$$\rho(x) = \frac{1}{N} \sum_{n=1}^{N} \left(\frac{f(T^{a_n}x)}{Af(x)} \right)^p.$$

We note that $\int (Af)^p \rho = \int f^p \le \left(\frac{2}{C}\right)^p \int (Af)^p$. This gives an upper bound on ρ on part of the space. If $\rho(x)$ is small, then we have $(1/N)\sum_{n=1}^N f(T^{a_n}x)/Af(x) = 1$ and $(1/N)\sum_{n=1}^N (f(T^{a_n}x)/Af(x))^p$ is small. This implies that for most n, $f(T^{a_n}x)$ is comparable to $Af(x)$.

We use the above inequality to estimate $\|Af - Bf\|$ in three parts.

$$\begin{aligned} \|Af - Bf\| \le & \left(\int_{\{x\colon \rho(x) > L\}} |Af(x)|^p \right)^{1/p} \\ & + \left(\int_{\{x\colon \rho(x) \le L\}} \left(\sum_k 1_{E^k}(x) \frac{1}{N} \sum_{\{n\colon f(T^{a_n}x) \le R^{k-1}} f(T^{a_n}x) \right)^p \right)^{1/p} \\ & + \left(\int_{\{x\colon \rho(x) \le L\}} \left(\sum_k 1_{E^k}(x) \frac{1}{N} \sum_{\{n\colon f(T^{a_n}x) > R^{k+1}} f(T^{a_n}x) \right)^p \right)^{1/p}. \end{aligned} \tag{3}$$

First, for the second part of (3), we note that for any x, x belongs to some E^k. We then calculate $\frac{1}{N}\sum_{\{n\colon f(T^{a_n}x) \le R^{k-1}\}} f(T^{a_n}x) < R^{k-1} \le Af(x)/\sqrt{R} \le Af(x)/8$. It follows that the contribution of the second term is dominated by $\|Af\|/8$.

For the first term, we have

$$\begin{aligned} \left(\frac{2}{C}\right)^p \int (Af)^p &\ge \int_{\{x\colon \rho(x) > L\}} \rho \cdot (Af)^p \\ &\ge L \int_{\{x\colon \rho(x) > L\}} (Af)^p, \end{aligned}$$

so that the first term is dominated by $(2/C)L^{-1/p}\|Af\| = \|Af\|/8$.

Finally, for the third term, let x satisfy $\rho(x) \le L$. Since the E_k partition the space, we let $x \in E_k$. We have

$$\begin{aligned} L &\ge \frac{1}{N} \sum_{\{n\colon f(T^{a_n}x) > R^{k+1}\}} \left(\frac{f(T^{a_n}x)}{Af(x)}\right)^p \\ &= \frac{1}{N} \sum_{\{n\colon f(T^{a_n}x) > R^{k+1}\}} \frac{f(T^{a_n}x)}{Af(x)} \left(\frac{f(T^{a_n}x)}{Af(x)}\right)^{p-1}. \end{aligned}$$

It follows that

$$\frac{1}{N} \sum_{\{n\colon f(T^{a_n}x) > R^{k+1}\}} f(T^{a_n}x) R^{(p-1)/2} \le LAf(x),$$

so that $\frac{1}{N}\sum_{\{n\colon f(T^{a_n}x)>R^{k+1}\}} f(T^{a_n}x) \le (L/R^{(p-1)/2})Af(x)$. We see that the contribution from the last term is dominated by $(L/R^{(p-1)/2})\|Af\| \le \|Af\|/8$.

To complete the proof, we note that $\|Bf - B'f\| \le \|Af/\sqrt{R}\| \le \|Af\|/8$ so that $\|Af - B'f\| \le \|Af\|/2$. □

Lemma 4.8. *Let $J' \subset J$ and suppose $(F_j)_{j\in J'}$ and $(G_j)_{j\in J'}$ satisfy $G_j \le F_j$ and $\|G_j - F_j\| \le (C/4)\|f\|_p$. Suppose further that $\|\max_{j\in J'} F_j\| \ge (C/2)|J|^{1/p}\|f\|$. Then*

$$\|\max_{j\in J'} G_j\| \ge (C/4)|J|^{1/p}\|f\|$$

Proof. Let $H = \max_{j\in J'} F_j(x) - \max_{j\in J'} G_j(x)$ and let E_j be the set $\{x\colon F_j(x) = \max_{k\in J'} F_k(x)\}$. Then

$$\begin{aligned} \|H\|^p = \int H^p &= \sum_{j\in J'} \int_{E_j} H^p \\ &= \sum_{j\in J'} \int_{E_j} (F_j - \max_{k\in J'} G_k)^p \\ &\le \sum_{j\in J'} \int_{E_j} (F_j - G_j)^p \\ &\le \sum_{j\in J'} (C/4)^p\|f\|^p \le (C/4)^p|J|\cdot\|f\|^p. \end{aligned}$$

We then have that $\|\max_{j\in J'} G_j\| \ge \|F\| - \|H\| \ge (C/4)|J|^{1/p}\|f\|$. □

We now assemble the above lemmas to complete the proof of Theorem 4.4.

Proof of Theorem 4.4: (4) implies (3) and (5). Recall that we are assuming that $\|\max_{j\in J} A_j f\| = C|J|^{1/p}\|f\|$. From Lemma 4.6, we can pick a subset J' of J such that for each $j \in J'$, $\|A_j f\| \ge (C/2)\|f\|$ and also such that $\|\max_{j\in J'} A_j f\| \ge (C/2)|J|^{1/p}\|f\|$.

By Lemma 4.7, there exists an R and operators B'_j for each $j \in J'$ such that $\|A_j f - B'_j f\| \le \|A_j f\|/2$. Specifically, these operators were defined as follows. Let $E_j^k = \{x\colon A_j f(x) \in (R^{k-1/2}, R^{k+1/2}]\}$ and define operators by

$$B_j^k f(x) = \frac{1}{|I_j|} \sum_{\{n\in I_j\colon f(T^{a_n}x)\in(R^{k-1},R^{k+1}]\}} f(T^{a_n}x)\mathbf{1}_{E_j^k}(x),$$

where I_j is the range of indices of the a_n involved in the jth average. We then define ${B'}_j^k f(x) = B_j^k f(x)\mathbf{1}_{\{x\colon B_j^k f(x) > R^{k-1}\}}$ and ${B'}_j f(x) = \sum {B'}_j^k f$. Note that the non-zero values taken by ${B'}_j^k$ are in the range $(R^{k-1}, R^{k+1}]$.

Applying Lemma 4.8 with $F_j = A_j f$, $G_j = B'_j f$, we deduce $\|\max_{j\in J'} B'_j f\| \geq (C/4)\|f\|$.

Now write f as a decomposition $f = \ldots + f_{-1} + f_0 + f_1 + f_2 + \ldots$, where $f_k(x) = f(x)\cdot \mathbf{1}_{\{x\colon R^{k-1}<f(x)\leq R^k\}}$. Suppose that $\max_{j\in J'} B'_j f(x) \in (R^{k-1}, R^k]$. We will assume that the maximum is attained for $\ell \in J'$. Since the maximum is in the range $(R^{k-1}, R^k]$, it follows that $B'_\ell f(x) = {B'}_\ell^k f(x)$ or ${B'}_\ell^{k-1} f(x)$. In particular, we have $B'_\ell f(x) = B'_\ell(f_{k-2} + f_{k-1} + f_k)(x)$. Setting $h_k = f_{k-2} + f_{k-1} + f_k$, if x satisfies $\max_{j\in J'} B'_j f(x) \in (R^{k-1}, R^k]$, we have shown that $\max_{j\in J'} B'_j f(x) = \max_{j\in J'} B'_j h_k(x)$.

Write $F(x) = \max_{j\in J'} B'_j f(x)$ and decompose F into parts F_k where $R^{k-1} < F(x) \leq R^k$. Then the above shows that $\max_j B'_j h_k \geq F_k$. We now have

$$\begin{aligned}
(C/12)^p |J| \sum_k \|h_k\|^p &\leq (C/4)^p |J| \sum_k \|f_k\|^p \\
&= (C/4)^p |J| \|f\|^p \\
&\leq \|F\|^p = \sum_k \|F_k\|^p \\
&\leq \sum_k \|\max_j B'_j h_k\|^p
\end{aligned}$$

In particular, there must exist a k such that $\|h_k\| \neq 0$ and $\|\max_{j\in J'} B'_j h_k\| \geq (C/12)|J|^{1/p}\|h_k\|$.

Since h_k only takes non-zero values between R^{k-3} and R^k, we see from the definition of B' that $\max_{j\in J'} B'_j h_k$ only takes non-zero values between R^{k-4} and R^k. It follows that $\|\max_{j\in J} A_j h_k\|_{p,\infty} \geq \|\max_{j\in J'} B'_j h_k\|_{p,\infty} \geq (C/12R^4)|J|^{1/p}\|h_k\|$.

Further, since both h_k and $\max_{j\in J'} B'_j h_k$ take on values in ranges with a bounded ratio between the endpoints, it follows that for any $1 < q < \infty$, there exists a C' such that $\|\max_{j\in J} A_j h_k\|_{q,\infty} \geq \|\max_{j\in J'} B'_j h_k\|_{q,\infty} \geq C'|J|^{1/q}\|h_k\|_q$. Applying the equivalence (3) implies (1) for L^q completes the proof of the theorem.

□

The proof of Theorem 2.4 will depend on the following lemmas.

Lemma 4.9. *Suppose the sequence (a_n) of real numbers satisfies the following condition:*

> *For each positive integer M, there exists an n_0 such that if $N \geq n_0$ and K satisfies $K \leq MN$ then for any sequence $r_1, r_2, \ldots, r_n$ of integers, there is a positive real α so that for*
>
> $$\lfloor \alpha a_{N+k} \rfloor \equiv r_k \mod K, \quad \textit{for all } 1 \leq k \leq n.$$

Then the sequence of times (a_n) is ultimately bad for L^1.

Lemma 4.10. *Suppose the sequence (a_n) of real numbers satisfies the following condition:*

There exists an n_0 such that if $N \geq n_0$ and K satisfies $K \leq N$ then for any sequence $r_1, r_2, \ldots, r_n$ of integers, there is a positive real α so that

$$\lfloor \alpha a_{2^{N-1}+k} \rfloor \equiv r_k \mod K, \quad \text{for all } 1 \leq k \leq n.$$

Then the sequence of times (a_n) is ultimately bad in every L^p ($p > 1$).

Corollary 4.11. *Suppose that for some fixed ϵ the sequence (a_n) satisfies*

$$\frac{a_{n+1}/a_n}{n^\epsilon} \to \infty.$$

Then (a_n) is ultimately bad in L^1.

Proof. This follows from Lemma 4.9 using a standard lacunarity argument. Let the sequence (a_n) be as in the statement of the lemma. Using Remark 4.2, we first refine to a subsequence (b_n) of the (a_n)'s occurring with bounded gaps in the original sequence such that $(b_{n+1}/b_n)/n^2 \to \infty$. It will be sufficient to show that (b_n) is ultimately bad for L^1. Let M be given. Then there exists an $N > 2M$ such that $n \geq N$ implies $b_{n+1}/b_n > n^2 > 2MN$. Let $K \leq MN$, we see that for any $n \geq N$, $b_{n+1}/b_n > 2K$.

To finish the argument, we claim that for any sequence $r_0, r_1, \ldots, r_{t-1}$ of integers with $0 \leq r_i < K$, there exists an interval I of length $1/(Kb_{t-1})$ such that for $\alpha \in I$, $\alpha b_i \bmod 1 \in [r_i/K, (r_i+1)/K)$ for $0 \leq i < t$.

We prove this by induction. Clearly it is true for $t = 1$. Suppose that it holds for $t \leq s$ and let I be the interval of length $1/(Kb_{s-1})$ such that for $\alpha \in I$, $\alpha b_i \bmod 1 \in [r_i/K, (r_i+1)/K)$ for $0 \leq i < s$. We see that $S = \{\beta\colon b_s\beta \bmod 1 \in [r_s/K, (r_s+1)/K)$ is a union of intervals of length $1/(Kb_s)$ spaced $1/b_s$ apart. Since $1/b_s(1/K+1) < 1/(Bb_{s-1})$ we see that I contains a complete interval from S. Letting J be the subinterval, the induction is complete. □

Corollary 4.12. *Suppose that for some fixed ϵ the sequence (a_n) satisfies*

$$\frac{a_{n+1}/a_n}{(\log n)^\epsilon} \to \infty.$$

Then (a_n) is ultimately bad in L^p for every $1 < p < \infty$.

Proof. The proof is similar to that of Corollary 4.11. Let (a_n) be as in the statement and suppose that $(a_{n+1}/a_n)/(\log n)^\epsilon \to \infty$. First, refine the sequence to a subsequence (b_n) with bounded gaps in the original sequence so that the new sequence (b_n) satisfies $(b_{n+1}/b_n)/(\log n) \to \infty$. There exists an n_0 such that $n \geq 2^{n_0-1}$ implies $b_{n+1}/b_n > 3\log(2n)$. Now if $K \leq N$, we see that for any $n > 2^{N-1}$, $b_{n+1}/b_n > 3\log(2n) > 2N \geq 2K$. The remainder of the argument follows exactly as in Corollary 4.11 □

Remark. Unfortunately both in the L^1 and L^p cases, an arbitrary lacunary sequence (a_n) is out of reach for now.

Corollary. *The sequence $(n!)$ is ultimately bad in every L^p, $1 \leq p < \infty$.*

Proof. This is an immediate consequence of Corollaries 4.11 and 4.12. □

Proof of Theorem 2.4. This follows from the above Corollary. □

Corollary. *Let the sequence* (a_n) *be independent over the rationals. Then* (a_n) *is ultimately bad in every* L^p, $1 \le p < \infty$.

Proof. This follows since the vectors $\alpha(a_N, a_{N+1}, \dots, a_{N+t-1}) \bmod 1$ are dense in the t-dimensional torus as α runs over the positive reals. Accordingly, for any N, K and $r_0, \dots, r_{t-1}$, there exists a positive real number α with the property that $\alpha(a_N, a_{N+1}, \dots, a_{N+t-1}) \bmod 1 \in \prod_{0 \le i < t}[r_i/K, (r_i+1)/K)$. ☐

Corollary. *The sequence* $(\sqrt{n})$ *is ultimately bad in every* L^p, $1 \le p < \infty$.

Proof. The set $\{\sqrt{s}\colon s \text{ squarefree}\}$ is independent over the rationals and arranged in increasing order it forms a positive density subsequence of $(\sqrt{n})$ so that there exists a fixed k such that the first 2^n squarefree numbers are a subset of the first 2^{n+k} square roots. We then use the fact that if (w_t) satisfies $C_p(w) = \infty$, then $C_p(w_{t+k}) = \infty$, so letting A_t be the average over the first 2^t square roots and B_t be the average over the first 2^t squarefree square roots, we have the estimate $w_{t+k}A_{t+k}f \ge 2^k w_{t+k}B_t f$. Since $C(w_{t+k}) = \infty$, the right hand side can be made to diverge and hence so does the left hand side. ☐

Proof of Lemma 4.9. We just deny the maximal inequality on $[0,1)$. Let the positive integer M be given and let n_0 be as in the statement of the lemma. We consider the set of N such that

$$\sum_t [1/N < w_t < 2^t/(4N)]w_t > 3M. \tag{4}$$

We note that these terms are necessarily unbounded above as $\sum_t [z < w_t < 2^t z]w_t$ may be bounded above by the sum of two of these terms. We assert that the set of N satisfying (4) is unbounded above. This is because either $\limsup w_t > 0$, in which case for all large enough N we have $\sum_t [1/N < w_t < 2^t/N]w_t = \infty$, or $w_t \to 0$, in which case for $N < K$, $\sum_t [1/N < w_t < 2^t/N]w_t$ is uniformly bounded above by $\sum_t [w_t > 1/K]w_t$ which is the sum of a finite number of terms and hence is finite. Since the sums $\sum_t [1/N < w_t < 2^t/(4N)]w_t$ are as noted above unbounded in N, there must exist arbitrarily large integers N for which the sum exceeds $3M$. Hence we may choose an N satisfying (4) such that $N > n_0$.

If we consider the t's such that $2^t \le 2N$, then we see $\sum_{\{t\colon 2^t \le 2N\}}[1/N < w_t < 2^t/(4N)]w_t < \sum_{\{t\colon 2^t \le 2N\}} 2^t/(4N) < 1$ so that $\sum_{\{t\colon 2^t > 2N\}}[1/N < w_t < 2^t/(4N)]w_t > 2M$. It then follows that there is a finite set U of the t's satisfying $2^t > 2N$ and $1/N < w_t < 2^t/(4N)$ so that we still have $\sum_{t \in U} w_t > 2M$.

Now set $K = NM$. For each $t \in U$, select $\lfloor Nw_t \rfloor$ different residue classes modulo K. Denote these residue classes by R_t. Since

$$\sum_{t \in U} \lfloor Nw_t \rfloor > NM,$$

we can choose the R_t so that their union over $t \in U$ covers all residue classes modulo K. Since $2^{t-1} > N$ for all $t \in U$, we can now apply the condition in the statement of the lemma to conclude that there is a positive α so that for each $t \in U$ and $r \in R_t$ there are at least $2^{t-1}/(Nw_t)$ n's between 2^{t-1} and 2^t with $\lfloor \alpha a_n \rfloor \equiv r \bmod K$. Define the function f by

$$f(x) = \begin{cases} 2N, & \text{if } 0 \le x < \frac{2}{K} \\ 0, & \text{otherwise} \end{cases}$$

Clearly, $\|f\|_1 = \frac{4}{M}$.

We define a measure-preserving flow on $[0,1]$ by $T^\zeta(x) = x + \alpha\zeta/K$. Now let $x \in [0,1)$ be arbitrary. Then there is a t and $r \in R_t$ with $x \in [\frac{K-r}{K}, \frac{K-r+1}{K})$. Since we have a number of n between 2^{t-1} and 2^t such that $\lfloor \alpha a_n \rfloor \equiv r \mod K$, for these n, we see that $T^{a_n}(x) \in [0, 2/K)$ so that $f(T^{a_n}(x)) = 2N$ and hence we can estimate:

$$\begin{aligned}
&w_t \frac{1}{2^t} \sum_{n \le 2^t} f(x + a_n) \\
&\ge w_t \frac{1}{2^t} \sum_{\{n\colon 2^{t-1} < n \le 2^t\}} \big[\lfloor \alpha a_n \rfloor \equiv r \mod K\big] f(T^{a_n}(x)) \\
&= w_t \frac{1}{2^t} \sum_{\{n\colon 2^{t-1} < n \le 2^t\}} \big[\lfloor \alpha a_n \rfloor \equiv r \mod K\big] 2N \\
&\ge w_t \frac{1}{2^t} 2N \frac{2^{t-1}}{N w_t} \\
&= 1.
\end{aligned}$$

We have shown that

$$m\left\{x\colon \sup_t w_t \frac{1}{2^t} \sum_{n \le 2^t} f(T^{a_n}(x)) \ge 1\right\} = 1 > \frac{M}{4}\|f\|_1.$$

Since M is arbitrary, the required violation of the maximal inequality is shown.

□

Proof of Lemma 4.10. We aim to establish condition (3) of Theorem 4.4 for the sequence (a_n). Let n_0 be as in the statement of the lemma and let J be a finite subset of $\mathbb{N}$. If $|J| \le 2n_0$, taking f to be a constant function, we have $\|\max_{j\in J} A_j f\|_{p,\infty} \ge (2n_0)^{-1/p}|J|^{1/p}\|f\|_p$.

If on the other hand, $|J| \ge 2n_0$, then let $N = K = \lfloor |J|/2 \rfloor$ and let $J' = \{j_1, \ldots, j_K\}$ be a subset of J of size K consisting of elements of J at least as big as $|J|/2$. By assumption, there exists an α such that for $n \in \{2^{j_\ell - 1} + 1, 2^{j_\ell - 1} + 2, \ldots, 2^{j_\ell}\}$, $\lfloor \alpha a_n \rfloor \equiv \ell - 1 \pmod K$. Then letting (T^t) be the flow on $[0, K)$ given by $T^t(x) = x - \alpha t$ and f be the function $2 \cdot \mathbf{1}_{[0,2)}$, we see that $\|f\|_p = 2^{1+1/p}|K|^{-1/p}$. We also have for $x \in [n, n+1)$, $A_{j_n} f(x) \ge 1$. It follows that $\|\max_{j\in J} A_j f(x)\|_{p,\infty} = 1 \ge 2^{-1-2/p}|J|^{1/p}\|f\|$ as required.

□

5. Khintchine's Conjecture

In this section, we consider the averages arising in a conjecture due to Khintchine. For $f \in L^p[0,1)$, Write $T_n f(x) = f(nx \bmod 1)$. Khintchine [8] conjectured in 1923 that for every $f \in L^1$, it is the case that $1/N \sum_{n=1}^N T_n f(x)$ converges pointwise almost everywhere to $\int f$. This was answered negatively by Marstrand [12] in 1970. This negative result was strengthened further in Bourgain's work using his Entropy Method [4].

We start with a lemma showing the equivalence of maximal theorems for averages of the type $\frac{1}{N}\sum_{n\le N} f(a_n x)$ for functions $f \in L^p([0,1))$ and averages of the type $\frac{1}{N}\sum_{n\le N} g(x - \log a_n)$ for functions $g \in L^p(\mathbb{R})$.

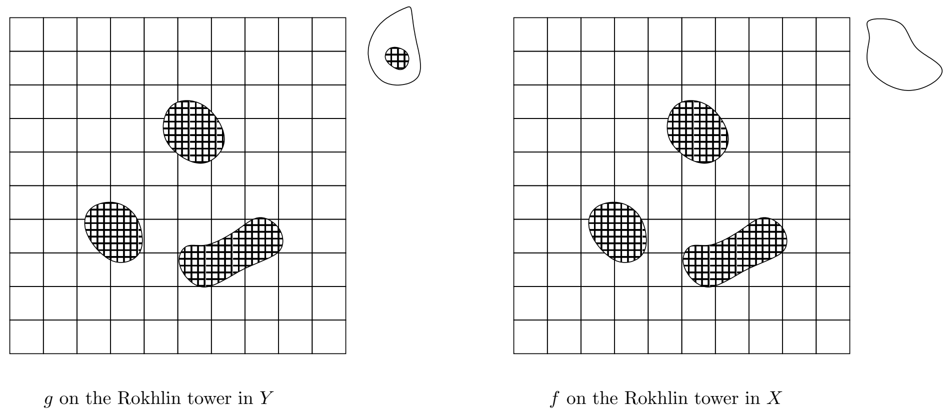

FIGURE 1. Copying a function between Rokhlin towers

This lemma will allow us to give a very simple demonstration of Marstrand's result and in fact to show more: that the sequence of operators (T_n) is ultimately bad in L^p for $p > 1$.

We also take up a question posed by Nair in [13] concerning a version of Khintchine's conjecture, where the $T_n f(x)$ are averaged along a subsequence rather than all of the integers. Nair proved that if the sequence (a_n) is the increasing enumeration of a finitely generated multiplicative subsemigroup of the positive integers, then for all $f \in L^1$, the averages $1/N \sum_{n=1}^{N} T_{a_n} f(x)$ converge for almost every x to $\int f$. He asked about the case of averaging along an infinitely generated subsemigroup of the positive integers.

Later, Lacroix [10, 11] took up this question and claimed that there do exist infinitely generated subsemigroups of the integers along which the above averages converge. Unfortunately, while the arguments in his papers appear to be correct, the result seems to be false as they rely on an incorrect statement in Krengel's book [9].

Here, using the lemma again, we clear up the situation with an explicit dichotomy in Theorem 5.7. If S is a multiplicative subsemigroup of the positive integers, then the averages above converge for all $f \in L^1$ if and only if S is contained in a finitely generated subgroup of the positive integers.

Lemma 5.1. *Let (a_n) be any sequence of positive integers. Let $I_1, \ldots, I_k$ be any non-empty finite subsets of $\mathbb{N}$. Denote by $A_j f(x)$ the average $1/|I_j| \sum_{n \in I_j} f(a_n x)$ and by $B_j g(y)$ the average $1/|I_j| \sum_{n \in I_j} g(y - \log a_n)$. Then the following are equivalent:*

(1) *There exists an $f \in L^p[0,1)$ such that $\| \max_{j \le k} A_j f \|_{p,\infty} > C \|f\|_p$;*
(2) *There exists a $g \in L^p(\mathbb{R})$ such that $\| \max_{j \le k} B_j g \|_{p,\infty} > C \|g\|_p$.*

Let $T_n(x) = nx \bmod 1$ and let $S_n(y) = y - \log n$. The crux of the proof is the simple observation that S_n and T_n satisfy the same basic relationship: $S_{nm} = S_n \circ S_m$ and $T_{nm} = T_n \circ T_m$, allowing data from a Rokhlin tower for one system to be copied to a Rokhlin tower for the other. This transference is illustrated in Figure 1.

Proof. Let $p_1, p_2, \dots, p_d$ be the primes occurring in the prime factorization of elements of $\{a_n : n \in \bigcup I_j\}$ and let r be the maximum of all the powers of the p_d occurring in the elements of $\{a_n : n \in \bigcup I_j\}$.

We first observe that condition (2) is equivalent to the following condition that we call (2'):

> There exists an M rationally independent of $\{\log p_1, \dots, \log p_d\}$ and a $g \in L^p([0, M))$ such that $\| \max_{j\le k} B_j g\|_{p,\infty} > C\|g\|_p$, where the difference $y - \log a_n$ is interpreted modulo M.

To see that (2) implies (2'), simply restrict the function g occurring in (2) to some large interval, whereas to see that (2') implies (2), starting from the function in (2'), concatenate a large number of translated copies of the function g on intervals $[(n-1)M, nM)$ to produce a function supported on $[0, LM)$ and observe that condition (2) is satisfied.

We will therefore demonstrate the equivalence of (1) and (2'). If (2') holds, let M be as in the statement, otherwise let $M = 1$ so that M is rationally independent of $\{\log p_1, \dots, \log p_d\}$. Let N be chosen to be a large integer and let $\epsilon > 0$ be small.

For $\mathbf{n} \in \mathbb{N}^d$, write $T^{\mathbf{n}}(x) = \prod p_i^{n_i} x \bmod 1$, and for $\mathbf{n} \in \mathbb{Z}^d$, write $S^{\mathbf{n}}(y) = y - \sum n_i \log p_i \bmod M$. We observe that these are both free actions. Accordingly, we can construct Rokhlin towers of geometry $\Lambda_N = \{0, 1, \dots, N-1\}^d$ for both systems with an error set of size exactly ϵ: there exist $V \subset [0, M)$ and $W \subset [0, 1)$ such that $\mu(V) = \lambda(W)$ and that the sets $S^{-\mathbf{n}}V$ for $\mathbf{n} \in \Lambda_N$ are mutually disjoint, as are the sets $T^{-\mathbf{n}}W$. We now construct measure-preserving maps between the Rokhlin towers. Let $R_X = \bigcup_{\mathbf{n}\in\Lambda_N} T^{-\mathbf{n}}W$ and $R_Y = \bigcup_{\mathbf{n}\in\Lambda_N} S^{-\mathbf{n}}V$.

Let θ_0 be an arbitrary measure-preserving measurable bijection from V to W. Then define $\theta(x)$ for $x \in T^{-\mathbf{n}}W$ by $S^{-\mathbf{n}} \circ \theta \circ T^{\mathbf{n}}(x)$. Similarly, letting $\mathbf{k} = (N-1, \dots, N-1)$, let ψ_0 be an arbitrary measure-preserving measurable bijection from $S^{-\mathbf{k}}V$ to $T^{-\mathbf{k}}W$ and define $\psi(y)$ for $y \in S^{-\mathbf{n}}V$ by $\psi(y) = T^{\mathbf{k}-\mathbf{n}} \circ \psi_0 \circ S^{-(\mathbf{k}-\mathbf{n})}(y)$. These are then defined so as to ensure that $\theta(T^{\mathbf{n}}(x)) = S^{\mathbf{n}}(\theta(x))$ provided that the orbit of x remains inside the tower and similarly $\psi(S^{\mathbf{n}}(y)) = T^{\mathbf{n}}(\psi(y))$.

If condition (2') holds, we define f on R_X by $f(x) = g(\theta(x))$ and define f to be 0 on the remainder of $[0, 1)$. By construction, we see that provided that $x \in \bigcup_{\{\mathbf{n}\colon r\le n_i<N, i=1,\dots,d\}} T^{-\mathbf{n}}(W)$, we have that $f(T^{\mathbf{n}}x) = g(S^{\mathbf{n}}(\theta(x))$ for $\mathbf{n}$ with coefficients less than r. In particular, since the times a_n involved in the averages A_j and B_j for $j \in J$ may be expressed in terms of $p_1, \dots, p_d$ with powers at most r, we see that for such an x, we have $\max_{j\in J} A_j f(x) = \max_{j\in J} B_j g(\theta(x))$. Now for sufficiently small ϵ and large N, we will have $\| \max_{j\le k} A_j f(x)\|_{p,\infty} > C\|f\|_p$.

If condition (1) holds, we define g on R_Y by $g(y) = f(\psi(y))$ and define g to be 0 elsewhere. The same argument as above demonstrates that condition (2') holds provided that N is chosen to be sufficiently large and ϵ is taken to be sufficiently small.

□

Theorem 5.2. *The sequence (T_n) of operators defined by $T_n f(x) = f(nx \bmod 1)$ is ultimately bad in L^p for all $p > 1$.*

Proof. We let $g(y)$ be the function $2 \cdot \mathbf{1}_{[0,2\log 2)}$ and set for any finite set $J \subset \mathbb{N}$, $I_j = \{n : n \le 2^j\}$. We will then demonstrate that $\max_{j\in J} \|B_j g\|_{p,\infty} \ge C|J|^{1/p}\|g\|_p$ for a constant C that does not depend on J. By Lemma 5.1, this will establish the existence of an $f \in L^1[0,1)$ satisfying condition (3) of Theorem 4.4 (see Remark 4.5).

We have $\|g\|_p^p = 2^{p+1}\log 2$. Let $j \in J$ and $x \in [j, j+1)$. Then $B_j(x) \geq 1$. It follows that the measure of the set where the maximal function exceeds 1 is at least $|J|$. This shows that $\| \max_{j\in J} B_t g(y)\|_{p,\infty} > C|J|^{1/p}\|g\|$, where $C = 2^{-1-1/p}(\log 2)^{1/p}$ as required. □

Lemma 5.3. *Let (h_n) be an non-decreasing sequence of real numbers and let $c_n = (h_n - h_{n-1})/h_n$ (or 0 in the case that the denominator is 0). Then $h_n = O(n^{\|c\|_{1,\infty}})$ as $n \to \infty$.*

Proof. We will suppose for simplicity that $h_1 > 0$. Suppose that $\|c\|_{1,\infty} = d < \infty$. We have $h_n = h_{n-1}/(1-c_n)$ so that in particular,

$$h_n = h_1 \prod_{j=2}^{n} \frac{1}{1-c_j}.$$

If (t_n) denotes the decreasing rearrangement of (c_n), then we have

$$h_n \leq h_1 \prod_{j=2}^{n} \frac{1}{1-t_j}.$$

Since $\|t\|_{1,\infty} = d$, we have $|\{j : t_j > d/k\}| < d/(d/k) = k$ so that $t_k \leq d/k$. Letting $C = h_1 \prod_{j=2}^{2d-1} 1/(1-t_j)$, we have

$$h_n \leq C \prod_{j=2d}^{n} \frac{1}{1-d/j}.$$

Taking logarithms, we see that $\log h(n) \leq C' + d\log n$ so that $h(n) \leq Kn^d$ as required. □

Theorem 5.4. *Let $(t_n)_{n\in\mathbb{N}}$ be an increasing sequence of real numbers with the property that $t_n \to \infty$. Let $h(N)$ denote $|\{n : t_n \leq N\}|$. If $\limsup h(N)/N^k = \infty$ for all k, then $B_N g$ fails to satisfy a maximal inequality, where $B_N g(y) = (1/h(N))\sum_{\{n:\, t_n \leq N\}} g(T_{t_n} y)$.*

Proof. Set $c_n = (h(n) - h(n-1))/h(n)$. For simplicity, we assume that $h(1) \geq 1$. We let g be the indicator function $\mathbf{1}_{[0, 2\log 2)}$ and we estimate $\|\sup_N B_N g\|$. We quickly see that for $n \leq x < n+1$, $B^* g(x) \geq B_n g(x) \geq (h(n) - h(n-1))/h(n) = c_n$. It follows that $\|\sup_N B_N g\|_{1,\infty} \geq \|c\|_{1,\infty}$.

By Lemma 5.3, since we know that for all k, $\limsup_{n\to\infty} h(n)/n^k = \infty$, it follows that $\|c\|_{1,\infty} = \infty$. □

The following corollary is closely related to a theorem of Jones and Wierdl [7] (the hypothesis and conclusion are both weakened).

Corollary 5.5. *If (a_n) is an increasing sequence of real numbers with the property that for all $\epsilon > 0$, $a_n \leq n^\epsilon$ for all sufficiently large n, then $B_N g$ fails to satisfy a maximal inequality.*

Proof. If $a_n \leq n^\epsilon$ for all $n \geq n_0$, then $h(n) \geq n^{1/\epsilon}$ for $n \geq n_0^\epsilon$. □

If S is an infinite subset of $\mathbb{N}$, we let S_N denote $\{n \in S : n \leq N\}$. For a function $f \in L^1([0,1))$, we consider the averages $A_N f(x) = 1/|S_N| \sum_{n\in S_N} f(nx)$.

Corollary 5.6. *Let S be an infinite subset of $\mathbb{N}$. If S has the property that*

$$\limsup_{N\to\infty} |S_N|/(\log N)^k = \infty \text{ for all } k,$$

then there exists $f \in L^1$ such that $\limsup A_N f(x) = \infty$ almost everywhere.

Proof. By Fact 2.1, the conclusion is equivalent to establishing the fact that there is no maximal inequality for the averages A_N. By Lemma 5.1, this is equivalent to establishing that there is no maximal inequality for the averages $B_N g(y) = (1/|S_N|)\sum_{t\in\log(S_N)} g(y-t)$. Since the number of elements of $\log(S)$ up to K is equal to the number of elements of S up to e^K, which by hypothesis is not bounded by any power of K, Theorem 5.4 gives the desired conclusion. □

Theorem 5.7. *If S is a multiplicative subsemigroup of the positive integers, then there is pointwise convergence of $A_N f(x)$ to $\int f$ for all $f \in L^1$ if and only if S is contained in a finitely generated semigroup.*

Proof. Suppose that $\log b_1, \dots, \log b_d$ are rationally independent. Let B be the largest of the b's. It follows that for any $n \geq 1$, there are at least Cn^d terms of $S \cap [1, B^n]$.

If S is not contained in any finitely generated semigroup, it follows that for any k, there exist elements $b_1, b_2, \dots, b_k$ of S whose logarithms are rationally independent so that the hypothesis of Corollary 5.6 is satisfied showing that there exists an $f \in L^1$ such that $\limsup A_N f(x)$ is infinite almost everywhere.

In the case where S is contained in a finitely generated semigroup, we make use of an ergodic theorem for amenable group actions due to Ornstein and Weiss [14]. It is sufficient to establish that the sets S_N defined above form a Følner sequence. For convenience, we use additive notation. Specifically, since by assumption, S is contained in a finitely generated semigroup of the positive integers, let the primes that appear as factors of elements of S be $p_1, \dots, p_k$. Given $n \in S$, write $n = p_1^{\alpha_1}\cdots p_k^{\alpha_k}$ and we will associate n with the vector $(\alpha_1, \dots, \alpha_k) \in \mathbb{Z}_+^k$. These vectors span a lattice in $\mathbb{Z}^k$ whose dimension we will call d. Let L_+ be the intersection of $\mathbb{Z}_+^k$ with the lattice spanned by the vectors in S. In this notation, S_N corresponds to $\{(\alpha_1, \dots, \alpha_k) \in S \colon \sum \alpha_i \log p_i \leq \log N\}$. Clearly the S_N are nested. It remains to establish the following two conditions.

$$\text{For all } n \in S,\ \lim_{N\to\infty} |(n+S_N) \triangle S_N|/|S_N| = 0 \tag{5}$$

$$\text{There exists an } M \text{ such that for all } N,\ |S_N - S_N| \leq M|S_N|. \tag{6}$$

The second of these is seen as follows: If $x \in S_N - S_N$, then x may be expressed as $(x_1, \dots, x_k) = (\alpha_1, \dots, \alpha_k) - (\beta_1, \dots, \beta_k)$, where $(\alpha_1, \dots, \alpha_k)$ and $(\beta_1, \dots, \beta_k)$ are in S_N. It follows that $\alpha_i \leq (\log N)/(\log p_i)$ so that $|x_i| \leq (\log N)/(\log p_i)$. Clearly the number of such elements x is bounded above by an expression of the form $C(\log N)^d$. On the other hand, by the argument at the start of the proof, there are at least $C'(\log N)^d$ elements in S_N so condition (6) holds.

To establish condition (5), let $x \in S$. We need to estimate the cardinality of $(S_N + x) \triangle S_N$. Clearly this is twice the cardinality of $(S_N + x) \setminus S_N$. This difference is contained in $\{(\alpha_1, \dots, \alpha_k) \in L_+ \colon \log N < \sum \alpha_i \log p_i \leq \log N + \sum x_i \log p_i\}$. To estimate this, we will use a crude estimate for the number $L(y)$ of lattice points in $\{(x_1, \dots, x_k) \in L_+ \colon \sum \alpha_i \log p_i \leq y\}$. Let V be the d-dimensional vector space spanned by S equipped with the inherited d-dimensional Lebesgue measure λ. Let F denote a convex fundamental domain for the lattice L inside the vector space V and let T denote the set $V \cap \{(\alpha_1, \dots, \alpha_k)\colon \sum \alpha_i \log p_i \le 1\}$. We claim that $L(y) = y^d \lambda(T)/\lambda(F) + o(y^d)$. To see this, note that if h is the diameter of F and letting $P_{\text{int}}(y)$ denote the set of lattice points in L_+ whose h-neighborhood lies within yT and let $P_{\text{ext}}(y)$ denote the set of lattice points whose h-neighborhood intersects yT. We now have $\text{Int}_{2h}(yT) \subset P_{\text{int}}(y)+F \subset yT \subset P_{\text{ext}}(y)+F \subset B_{2h}(yT)$ so that $|P_{\text{int}}(y)|\lambda(F) \le y^d\lambda(T) \le |P_{\text{ext}}(y)|\lambda(F)$. Clearly we also have $|P_{\text{int}}(y)| \le L(y) \le |P_{\text{ext}}(y)|$ so it follows that $|L(y) - y^d\lambda(T)/\lambda(F)| \le P_{\text{ext}}(y) - P_{\text{int}}(y) \le \lambda\big(B_h(yT) \setminus \text{Int}_h(yT)\big)$. Since this region is contained in the union of $d+1$ slabs each of which having bounded thickness and dimensions linearly dependent on y, this quantity is $O(y^{d-1})$. It now follows that $L(\log N + \sum x_i \log p_i) - L(\log N) = O((\log N)^{d-1})$ so that $\lim_{N\to\infty} |(S_N + x) \setminus S_N|/|S_N| = 0$ as required.

□

6. Ultimate Badness of Exponential sequences

Theorem 6.1. *For any $k \in \{2, 3, \dots\}$, the sequence (k^n) is ultimately bad in L^1.*

Proof. We deny a maximal inequality by carefully using the standard lacunarity trick for a rotation of the circle. Let (w_t) be a sequence such that $C_1(w) = \infty$. Let M be a large integer and fix a $y \in \mathbb{N}$ such that

$$m_y = \sum_t [2^{-y} < w_t < 2^{t-y}] w_t > M. \tag{7}$$

Let $n_0 = \lfloor 2y \log_k 2 \rfloor$ so that $k^{n_0} \approx 2^{2y}$. Throughout the proof, K will be used to denote various quantities that can be bounded above or below independently of y and M. (The bounds may however depend on k). Let f be the function on the circle taking the value 2^y on an interval of length $3/(k^{n_0}-1)$ starting at $-1/(k^{n_0}-1)$ and extending to $2/(k^{n_0}-1)$ and 0 elsewhere so that $\|f\|_1 \le K 2^{-y}$.

We now construct a number α such that letting T be the rotation of the unit circle by $-\alpha$ and computing the averages

$$A_t f(x) = w_t \frac{1}{2^t} \sum_{n \le 2^t} f(x - k^n \alpha),$$

the maximal function $f^*(x) = \sup A_t f(x)$ has weak L^1 norm greater than $K m_y \|f\|_1$.

Initially divide the circle into intervals of length $1/(k^{n_0}-1)$. These intervals have endpoints whose base k expansions are periodic with period dividing n_0. We label each interval by a string of n_0 symbols in $\{0, \dots, k-1\}$ that form the repeated block of the left endpoint so that if $B \in \{0, 1, \dots, k-1\}^{n_0}$ then I_B is the interval with left endpoint equal to $0.\overline{B}$ in the base k expansion. We shall consider only those intervals whose left endpoint's expansion has period exactly n_0. This excludes a negligible fraction of the intervals.

Consider a t satisfying $8n_0 2^{-y} < w_t < 2^{t-y}$. We say that an interval I_B is *infected* at time 2^t if

$$w_t \frac{1}{2^t} \sum_{2^{t-1} < n \le 2^t} f(x - k^n \alpha) > 1 \text{ for all } x \in I_B.$$

We will show how to choose the digits of α's base k expansion from the 2^{t-1} position to the 2^t position in order to bound below the number of new intervals infected at time 2^t. Summing these contributions over t, will give a lower bound for the maximal function as required.

We say that two words are (cyclically) equivalent if one is a cyclic permutation of the other. List representatives of all of the equivalence classes in some order as B_1, B_2, …. Suppose that by time 2^{t-1}, the intervals corresponding to $B_1, \ldots, B_j$ and their cyclic permutations are infected. We then define the binary expansion of α starting from the 2^{t-1}st digit to be concatenations of B_{j+1} until the intervals corresponding to the members of the equivalence class become infected. At this point, define digits of α to be concatenations of B_{j+2} etc. If all of the the equivalence classes are exhausted before the 2^tth digit of the binary expansion is defined, this will ensure that the constant in the maximal inequality exceeds $K2^y$ which will be sufficient as the y can be chosen to be arbitrarily large. We estimate the number of intervals that can be infected up to time 2^t as follows:

Let v_r be the sequence obtained by cyclically permuting B_{j+1} to the left r times and let J_r be the interval corresponding to v_r. We observe that if $n \geq 2^{t-1} + n_0$, we can write n as $2^{t-1} + jn_0 + r$ for some $j \geq 1$ and $0 \leq r < n_0$. In this case, for $x \in J_r$, we notice that $f(x - k^n\alpha) = 2^y$. In order to be infected, the sum needs to exceed $2^t/w_t$ so we see that this needs to be repeated $\lceil 2^{t-y}/w_t \rceil$ times. After this number of repetitions, α starts following the next B in a similar manner. The number of repetitions of B in each block is therefore bounded above by $2^{t-y+2}/w_t$ (the extra 1 being an overestimate coming from the fact that we have no control of the location of $x - k^n\alpha$ while $j = 0$). Since each repetition has length n_0, the length of the block is bounded above by $2^{t-y+2}n_0/w_t$. Since we have 2^{t-1} digits available to define, we are able to infect the intervals in at least $K\lfloor 2^{t-1}/(2^{t-y+2}n_0/w_t)\rfloor$ equivalence classes. Since we ensured that $w_t > 8n_0 2^{-y}$, we see that the quantity being rounded is greater than 1. As each equivalence class has n_0 members, we see that the number of intervals infected is given by $K2^y w_t$. The measure of the infected intervals then exceeds $K2^{-y}w_t \geq Kw_t\|f\|_1$. The constant in the maximal inequality therefore exceeds Kl_y where

$$l_y = \sum_t [8n_0 2^{-y} < w_t < 2^{t-y}]w_t$$

We complete the proof by demonstrating that $l_y + l_{2y} > m_y - 9$.

We have

$$l_y + l_{2y} \geq \sum_t [2^{-y} < w_t < 2^{t-2y} \text{ or } 8n_0 2^{-y} < w_t < 2^{t-y}]w_t.$$

This yields

$$\begin{aligned} m_y - (l_y + l_{2y}) &\leq \sum_t [2^{t-2y} \leq w_t \leq 8n_0 2^{-y}]w_t \\ &\leq 16y2^{-y}\#\{t \colon 2^{t-2y} \leq 8n_0 2^{-y}\} \\ &\leq 16y2^{-y}\#\{t \colon 2^t \leq 2^{2y}\} \leq 32y^2 2^{-y} < 9 \end{aligned}$$

as required. ☐

Theorem 6.2. *For $k \in \{2, 3, \ldots\}$, the sequence (k^n) is ultimately bad for L^p when $p > 1$.*

Proof. We will use condition (3) established in Theorem 4.3 for ultimate badness. We deal with the case $k > 2$. The fact that (2^n) is ultimately bad follows from the fact that (4^n) is ultimately bad using Remark 4.2.

For a given subset J of the positive integers, we construct a characteristic function $f = 1_B$ on $\mathbb{Z}$ such that $A_j f$ (the average over the jth dyadic block) takes a value of order 1 on a set of size approximately $|B|$, but that $A_j f$ and $A_{j'} f$ are disjointly supported for distinct $j, j' \in J$.

Let J be a finite set of integers. We will assume that J contains no two consecutive integers. For $j \in J$, let B_j denote $(k-1) \cdot k^{2^j} \cdot \{1, 2, 3, 4, \ldots, k^{2^{j+1}}\}$. Let C_j denote the truncated version $(k-1) \cdot k^{2^j} \cdot \{1, 2, 3, 4, \ldots, k^{2^{j+1}} - k^{2^j}\}$.

Let $B = \sum_{j \in J} B_j$, $C = \sum_{j \in J} C_j$; and let $f \in l^p$ be the characteristic function of B. By the requirement that J contains no two consecutive integers, it follows that each element of B may be expressed in only one way as the sum of elements of the B_j's. Note that $|C_j|/|B_j| = 1 - k^{-2^j}$ so that since $|C| = \prod_{j \in J} |C_j|$ and $|B| = \prod_{j \in J} |B_j|$, we have $|C| \geq |B|/2$.

Let $x \in \mathbb{Z}$ satisfy $x + k^{2^{j_0}} \in C$, for some $j_0 \in J$. Let $n = 2^{j_0}$. Let $m \in [2^{j_0}, 2^{j_0+1})$. We have $x + k^n = \sum_{j \in J} c_j$, where $c_j \in C_j$. In other words, $x + k^n = \sum_{j \in J \setminus \{j_0\}} c_j + (k-1) \cdot k^n a$, where $0 < a \leq k^{2n} - k^n$. We now have $x + k^m = x + k^n + (k^m - k^n)$. Then $k^m - k^n = k^n(k^{m-n} - 1) = (k-1) \cdot k^n((k^{m-n} - 1)/(k-1))$ so that $x + k^m = \sum_{j \in J - \{j_0\}} c_j + (k-1) \cdot k^n(a + (k^{m-n} - 1)/(k-1))$. Since $a \leq k^{2n} - k^n$ and $m - n < n$, it follows that $a + (k^{m-n} - 1)/(k-1) \leq k^{2n}$ ensuring that $x + k^m \in B$. This establishes that for $x \in C - k^{2^{j_0}}$, $A_{j_0} f(x) \geq 1/2$.

We now show that these sets are disjoint. Suppose that x lies in $C - k^{2^l}$ and $C - k^{2^m}$ for distinct $l > m \in J$. Then we see that $k^{2^l} - k^{2^m} \in B - B$. We show that this gives rise to a contradiction as follows. Note that $B - B = \sum_{j \in J} S_j$, where $S_j = (k-1) \cdot k^{2^j} \cdot \{t \colon |t| < k^{2^{j+1}}\}$. If $a \in B - B$ has its largest non-zero summand in the S_j block, then we see that $k^{2^{j+2}}/2 > (k-1)(k^{3 \cdot 2^j} + k^{3 \cdot 2^{j-2}} + \ldots) \geq |a| \geq (k-1)(k^{2^j} - k^{3 \cdot 2^{j-2}} - k^{3 \cdot 2^{j-4}} \ldots) > k^{2^j}$ If $j \geq l$, we see that a exceeds $k^{2^l} - k^{2^m}$. If $j < l$ then $j \leq l - 2$ and we see that a is smaller than $k^{2^l} - k^{2^m}$. Note that this is where we made use of the assumption that $k > 2$.

It follows that $\| \sup_{j \in J} A_j f \|_{p,\infty} \geq (|J||C|)^{1/p}/2 \geq |J|^{1/p} \|f\|_p/4$. A standard argument using Rokhlin towers similar to (but simpler than) Lemma 5.1 allows this example to be transferred to an arbitrary aperiodic system. □

7. Questions and Remarks

Remark 7.1. The sequence of times $a_n = \lfloor \log n \rfloor$ is ultimately bad in L^p for all $1 < p < \infty$, but not in L^1. To see that the sequence is ultimately bad in L^p, using Theorem 4.4, we verify that for $f(x) = 2 \cdot \mathbf{1}_{[0, 2\log 2]}$ and $J \subset \mathbb{N}$, for $x \in [j, j+1]$, we see that $A_j f(x) \geq 1$, verifying condition (3) of Theorem 4.4.

To see that the sequence is not ultimately bad in L^1, let $w_t = 1/t$ and note that $C_1(w_t) = \infty$ but $\sup_t w_t A_t f$ is bounded above by the regular ergodic maximal function of f, which has weak L^1 norm bounded above by $\|f\|_1$.

Question 7.2. *Is the sequence of operators $M_n f(x) = f(nx \bmod 1)$ ultimately bad in L^1?*

Remark 7.3. We remark that Theorem 5.2 also shows that if $s(n)/\log n \to 0$, then there exists an $f \in L^1$ for which $1/(ns(n)) \sum_{k \leq n} f(kx)$ diverges almost everywhere. We also pose the following weakening of Question 7.2.

Question 7.4. *Does there exist $f \in L^1([0,1))$, such that $1/(n \log n) \sum_{k \leq n} f(kx)$ diverges almost everywhere?*

Question 7.5. *Do there exist lacunary sequences that are not ultimately bad in some L^p?*

Question 7.6. *If the sequence of times (a_n) is ultimately bad for L^1, does it follow that it is ultimately bad for L^p $(p > 1)$. Remark 7.1 shows that the converse is false.*

References

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

[2] M. Akcoglu, R. Jones, and J. Rosenblatt. The worst sums in ergodic theory. *Michigan Math. J.*, 47:265–285, 2000.

[3] C. Bennett and R. Sharpley. *Interpolation of operators*. Academic Press, 1988.

[4] J. Bourgain. Almost sure convergence and bounded entropy. *Israel J. Math.*, 63:79–97, 1988.

[5] C. Demeter. The best constants associated with some weak maximal inequalities in ergodic theory. *preprint*, 2003.

[6] A. Garsia. *Topics in almost everywhere convergence*. markham, 1970.

[7] R. Jones and M. Wierdl. Convergence and divergence of ergodic averages. *Ergodic Theory Dynam. Systems*, 14:515–535, 1994.

[8] A. Khintchine. Ein Satz über Kettenbrüche mit arithmetischen Anwendungen. *Math. Z.*, 18:289–306, 1923.

[9] U. Krengel. *Ergodic Theorems*. de Gruyter, 1985.

[10] Y. Lacroix. On strong uniform distribution II. *Acta Arith.*, pages 279–290, 1998.

[11] Y. Lacroix. On strong uniform distribution III. *Monatsh. Math.* to appear.

[12] J. M. Marstrand. On Khintchine's conjecture about strong uniform distribution. *J. London Math. Soc.*, 21:540–556, 1970.

[13] R. Nair. On strong uniform distribution. *Acta Arith.*, 56:183–193, 1990.

[14] D. Ornstein and B. Weiss. The Shannon–McMillan-Breiman theorem for a class of amenable groups. *Israel J. Math*, 44:53–60, 1983.

[15] J. Rosenblatt and M. Wierdl. Pointwise ergodic theorems via harmonic analysis. In K. Petersen and I. Salama, editors, *Ergodic theory and its connections with harmonic analysis*. Cambridge, 1995.

[16] S. Sawyer. Maximal inequalities of weak type. *Ann. Math.*, 84:157–174, 1966.

[17] E. M. Stein and N. J. Weiss. On the convegence of Poisson integrals. *Trans. Amer. Math. Soc.*, 140:35–53, 1969.

Abstract. We first study the rate of growth of ergodic sums along a sequence (a_n) of times: $S_N f(x) = \sum_{n \le N} f(T^{a_n} x)$. We characterize the maximal rate of growth and identify a number of sequences such as $a_n = 2^n$, along which the maximal rate of growth is achieved.

To point out though the general character of our techniques, we then turn to Khintchine's Strong Uniform Distribution Conjecture that the averages $(1/N)\sum_{n \le N} f(nx \bmod 1)$ converge pointwise to $\int f$ for integrable functions f. We give a simple, intuitive counterexample and prove that, in fact, divergence occurs at the maximal rate.

Department of Mathematical Sciences, University of Memphis, Memphis, TN 38152, USA

E-mail address: aquas@memphis.edu

E-mail address: mw@csi.hu