COLLOQUIUM MATHEMATICUM

VOL. 124 2011 NO. 2

POINTWISE CONVERGENCE FOR SUBSEQUENCES OF WEIGHTED AVERAGES

BY

PATRICK LAVICTOIRE (Madison, WI)

Abstract. We prove that if μ_n are probability measures on $\mathbb{Z}$ such that $\hat{\mu}_n$ converges to 0 uniformly on every compact subset of $(0,1)$, then there exists a subsequence $\{n_k\}$ such that the weighted ergodic averages corresponding to μ_{n_k} satisfy a pointwise ergodic theorem in L^1. We further discuss the relationship between Fourier decay and pointwise ergodic theorems for subsequences, considering in particular the averages along $n^2+\lfloor\rho(n)\rfloor$ for a slowly growing function ρ. Under some monotonicity assumptions, the rate of growth of $\rho'(x)$ determines the existence of a "good" subsequence of these averages.

1. Introduction. Generally speaking, if we have a family of operators T_n on a Banach space V which converge in some weak sense, we might ask whether there exists a subsequence T_{n_k} which converges in some stronger sense. An important special case here is the contrast between various types of "convergence in the mean" and "convergence almost everywhere", as for example in the following recent result of Kostyukovsky and Olevskii [8] on approximate identities.

Definition 1.1. A sequence of functions $\phi_n \in L^1(\mathbb{R})$ is an *approximate identity* on $\mathbb{R}$ if $\|\phi_n * f - f\|_1 \to 0$ as $n\to\infty$, for all $f\in L^1(\mathbb{R})$.

Theorem 1.2. *Let $\{\phi_n\}$ be an approximate identity on $\mathbb{R}$ consisting of non-negative functions. Then there is a sequence $\{n_k\}$ such that $\phi_{n_k} * f \to f$ a.e. for every $f\in L^1(\mathbb{R})$.*

As noted by Rosenblatt [10], this example is analogous to an open question about the pointwise convergence of subsequences of certain weighted ergodic averages. In that context, the natural analogue to an approximate identity is a sequence of probability measures $\{\mu_n\}$ such that for any ergodic dynamical system $(X,\mathcal{F},m,\tau)$ and any $f\in L^1$, the weighted averages

$$\mu_n f(x) := \sum_{j\in\mathbb{Z}} f(\tau^j x)\mu_n(j) \tag{1.1}$$

converge in the L^1 norm to $\int_X f\,dm$. This is equivalent [1, Proposition 1.7b

2010 *Mathematics Subject Classification*: Primary 42B25; Secondary 37A30.
Key words and phrases: subsequence ergodic theorem.

DOI: 10.4064/cm124-2-2

and Corollary 1.8], to the Fourier condition $\hat{\mu}_n(\gamma) \to 0$ for all $\gamma \in (0,1)$. However, a stronger condition seems to be required for an analogous result:

DEFINITION 1.3. We say that a sequence $\{\mu_n\}$ of probability measures on $\mathbb{Z}$ has *asymptotically trivial transforms* if $\hat{\mu}_n$ converges to 0 uniformly on every compact subset of $(0,1)$, or equivalently, if

$$\sup_{\gamma \in [0,1)} |(1 - e(\gamma))\hat{\mu}_n(\gamma)| \to 0,$$

where we denote $e(\gamma) = e^{2\pi i\gamma}$.

Bellow, Jones and Rosenblatt [1] proved the following:

THEOREM 1.4. *Suppose $\{\mu_n\}$ is a sequence of probability measures on $\mathbb{Z}$ with asymptotically trivial transforms. Then there exists a subsequence $\{n_k\}$ such that $\mu_{n_k} f(x)$ converges a.e. for every dynamical system $(X, \mathcal{F}, m, \tau)$ and every $f \in L^p$, $p > 1$.*

This was proved by an analysis of square functions and by interpolation from L^2 to L^p, which left open the L^1 question (see Section 4 in [1]). In this paper, I prove the following weak-type (1,1) maximal inequality on $\mathbb{Z}$:

THEOREM 1.5. *Suppose $\{\mu_n\}$ has asymptotically trivial transforms. Then there is a subsequence $\{n_k\}$ which obeys the weak type maximal inequality*

$$|\{x : \sup_k |\varphi * \mu_{n_k}(x)| > \lambda\}| \le C\lambda^{-1}\|\varphi\|_{\ell^1(\mathbb{Z})} \quad \forall \varphi \in \ell^1(\mathbb{Z}).$$

Given Theorem 1.4 and the Conze principle [6], this implies the full L^1 result:

COROLLARY 1.6. *Suppose $\{\mu_n\}$ has asymptotically trivial transforms. Then there exists a subsequence $\{n_k\}$ such that $\mu_{n_k} f(x)$ converges a.e. for every dynamical system $(X, \mathcal{F}, m, \tau)$ and every $f \in L^1(X)$.*

The next question is whether our stronger hypothesis (that $\{\mu_n\}$ has asymptotically trivial transforms) can be replaced by a weaker one (that $\hat{\mu}_n(\gamma) \to 0$ for all $\gamma \in (0,1)$). We strongly suspect that this cannot be done, and conjecture the following:

CONJECTURE 1.7. *There exists a sequence of probability measures $\{\mu_n\}$ such that $\hat{\mu}_n(\gamma) \to 0$ for all $\gamma \in (0,1)$, but for any subsequence $\{n_k\}$ and any (non–atomic) ergodic dynamical system $(X, \mathcal{F}, m, \tau)$, there exists an $f \in L^1(X)$ such that $\mu_{n_k} f(x)$ diverges on a set of positive measure in X.*

Finally, we examine a special case: the averages along the sequence $a_k := k^2 + \lfloor \rho(k) \rfloor$, where ρ is slowly growing. Such sequences (with k^2 replaced by an arbitrary polynomial) are of independent interest in the realm of pointwise ergodic theorems; Boshernitzan, Kolesnik, Quas and Wierdl [2] proved that within a broad class of integer sequences (subject to certain growth and regularity conditions), these are the only ones whose ergodic averages may diverge pointwise for some L^2 function, and they proved some necessary bounds and sufficient bounds on the growth of ρ for this to occur.

When we consider ergodic averages of L^1 functions and ask whether there exists a fixed subsequence of these that always converges, the positive result obtained via Corollary 1.6 and a negative result obtained by modifying a previous result of the author [9] meet at an exact threshold:

THEOREM 1.8. *Let $\rho \in C^2[0,\infty)$, with $\rho(x) \nearrow \infty$, $\rho'(x) \searrow 0$ and $\rho''(x) \nearrow 0$ as $x \to \infty$, be such that for some $\epsilon > 0$, $\rho'(x) \lesssim x^{-(\epsilon+2/3)}$ as $x \to \infty$. Consider the sequence of measures*

$$\mu_N := \frac{1}{N}\sum_{k=1}^{N} \delta_{k^2+\lfloor\rho(k)\rfloor}.$$

If $\rho'(x) \gg x^{-1}$ (thus $\rho(x) \gg \log x$), then the $\{\mu_N\}$ have asymptotically trivial transforms, and thus there exists a subsequence μ_{N_k} such that $\mu_{n_k} f(x)$ converges a.e. for every dynamical system $(X,\mathcal{F},m,\tau)$ and every $f \in L^1(X)$.

If $\rho'(x) \lesssim x^{-1}$ (thus $\rho(x) \lesssim \log x$), then for any subsequence $\{n_k\}$ and any (non-atomic) ergodic dynamical system $(X,\mathcal{F},m,\tau)$, there exists an $f \in L^1(X)$ such that $\mu_{n_k} f(x)$ diverges on a set of positive measure in X.

REMARK 1.9. This requires an additional monotonicity assumption (the existence of $\lim_{x\to\infty} x\rho'(x) \in [0,\infty]$) in order to become a true dichotomy; such an assumption is analogous to the Hardy field condition in [2].

REMARK 1.10. The condition $\rho'(x) \lesssim x^{-(\epsilon+2/3)}$ is an artifact of the proof rather than a genuine restriction. The requirement that $\rho'' \nearrow 0$, though, is necessary in some form to establish asymptotically trivial transforms; it is simple otherwise to create examples such that the exponential sums in (3.4) do not settle down away from 0.

2. Positive result for asymptotically trivial transforms. The proof of Theorem 1.5 makes use of the following technique: given the Calderón–Zygmund decomposition of a function $f = g+\sum_s b_s$, we classify s as "small", "large" or "intermediate" with respect to each term of our subsequence μ_n. For s "large", we can use a covering lemma to handle the terms; for s "small", we will use cancellation properties of b_s; and since for each s there will be only one n for which it counts as "intermediate", we can handle these terms with a trivial L^1 estimate. This idea plays a role in [12] as well as other papers.

The proof will also use a technique in singular integral theory, developed by Fefferman [7] and Christ [4] and first applied to ergodic theory by Urban and Zienkiewicz [13], which uses a sufficiently powerful L^2 estimate to prove a weak L^1 estimate.

We may assume that $\|\mu_n\|_{\ell^1(\mathbb{Z})} \le 1$ for all n, and write $\mu_n = \mu_n' + \eta_n$, where μ_n' is compactly supported and $\sum_{n=1}^{\infty} \|\eta_n\|_{\ell^1(\mathbb{Z})} < \infty$. Then

$$\begin{aligned}|\{x : \sup_n |\varphi * \eta_n(x)| > \lambda\}| &\le \lambda^{-1} \sum_{n=1}^{\infty} \|\varphi * \eta_n\|_{\ell^1(\mathbb{Z})} \\ &\le \lambda^{-1} \Big(\sum_{n=1}^{\infty} \|\eta_n\|_{\ell^1(\mathbb{Z})}\Big) \|\varphi\|_{\ell^1(\mathbb{Z})}.\end{aligned}$$

Now $|\mu_n'(\gamma)| \le |\hat{\mu}_n(\gamma)| + 2\|\eta_n\|_1$, and so $\hat{\mu}_n'$ converges to 0 uniformly on every compact subset of $(0,1)$. Thus we may assume that μ_n is compactly supported for each n.

Furthermore, if the union of these supports were compact, then it is easy to see (by Parseval's theorem) that $\|\mu_n\|_{\ell^1(\mathbb{Z})} \to 0$ and we may choose a subsequence such that $\sum_k \|\mu_{n_k}\|_{\ell^1(\mathbb{Z})} < \infty$; such a subsequence would trivially satisfy a weak maximal inequality.

We may therefore assume that the union of the supports of the μ_n is unbounded, and accordingly set $S(n) := \min\{s \ge 0 : \operatorname{supp} \mu_m \subset [-2^s, 2^s]$ for all $m \le n\}$, and $N(s) := \min\{n : S(n) > s\}$.

Since we will want the cancellation properties of μ_{n+1} to overcome the size of the support of μ_n, we choose an increasing subsequence $\{n_k\}$ such that

$$\sup_{\gamma \in [0,1)} |(1 - e(\gamma))\hat{\mu}_{n_k}(\gamma)| \le 2^{-2S(n_{k-1}) - 2k}$$

and such that $S(n_k)$ is strictly increasing. By passing to this subsequence, we may without loss of generality assume that μ_n has the following properties in the first place:

$$\operatorname{supp} \mu_n \subset [-2^{S(n)}, 2^{S(n)}], \tag{2.1}$$

$$\sup_{\gamma \in [0,1)} |(1 - e(\gamma))\hat{\mu}_n(\gamma)| \le 2^{-2S(n-1) - 2n}. \tag{2.2}$$

Now, given $\varphi \in \ell^1$ and $\lambda > 0$, we perform the discrete Calderón–Zygmund decomposition: we obtain a collection $\mathcal{B}$ of dyadic discrete intervals $Q_{s,k}$, and a decomposition $\varphi = g + \sum_{(s,k) \in \mathcal{B}} b_{s,k}$ with $\|g\|_\infty \le \lambda$, such that for all $(s,k) \in \mathcal{B}$,

$$\operatorname{supp} b_{s,k} \subset Q_{s,k}, \qquad \sum_x b_{s,k}(x) = 0, \qquad \sum_x |b_{s,k}(x)| \le \lambda |Q_{s,k}| = 2^s \lambda$$

and such that

$$\sum_{(s,k) \in \mathcal{B}} |Q_{s,k}| \le \lambda^{-1} \|\varphi\|_1.$$

Let $b_s := \sum_k b_{s,k}$ for each s, and let $Q^\star_{s,k}$ denote the interval with the same center as $Q_{s,k}$ and 3 times the length. Then $|\{x : \sup_n |\mu_n * \varphi(x)| > 3\lambda\}|$ is bounded by

$$\begin{aligned}|\{x : \sup_n |\mu_n * g(x)| > \lambda\}| &+ |\{x : \sup_n |\mu_n * b(x)| > 2\lambda\}| \\ &\leq 0 + \sum_{s,k} |Q^\star_{s,k}| + \Big|\Big\{x \notin \bigcup_{s,k} Q^\star_{s,k} : \sup_n |\mu_n * b(x)| > 2\lambda\Big\}\Big| \\ &\leq \frac{C}{\lambda}\|\varphi\|_1 + \Big|\Big\{x : \sup_n \Big|\mu_n * \sum_{s<S(n)} b_s(x)\Big| > 2\lambda\Big\}\Big|,\end{aligned}$$

because $\|\mu_n * g\|_\infty \leq \|\mu_n\|_1 \|g\|_\infty \leq \lambda$ and because $s \geq S(n) \Rightarrow \operatorname{supp}(\mu_n * b_{s,k}) \subset Q^\star_{s,k}$. Now $S(n-1) \leq s < S(n) \Rightarrow n = N(s)$, and therefore we decompose

$$\begin{aligned}\sup_n \Big|\mu_n * \sum_{s<S(n)} b_s(x)\Big| &\leq \sum_n \Big|\mu_n * \sum_{s=S(n-1)}^{S(n)-1} b_s(x)\Big| + \sup_n \Big|\mu_n * \sum_{s<S(n-1)} b_s(x)\Big| \\ &\leq \sum_s |\mu_{N(s)} * b_s(x)| + \sup_n \Big|\mu_n * \sum_{s<S(n-1)} b_s(x)\Big|.\end{aligned}$$

As mentioned earlier, we can trivially bound the contribution from the "intermediate" terms:

$$\begin{aligned}\Big|\Big\{x : \sum_s |\mu_{N(s)} * b_s(x)| > \lambda\Big\}\Big| &\leq \lambda^{-1} \sum_s \|b_s * \mu_{N(s)}\|_1 \\ &\leq \lambda^{-1} \sum_s \|b_s\|_1 \|\mu_{N(s)}\|_1 \leq \frac{C}{\lambda}\|\varphi\|_1.\end{aligned}$$

We have thus reduced this problem to the following claim:

Lemma 2.1.

$$\Big|\Big\{x : \sup_n \Big|\mu_n * \sum_{s<S(n-1)} b_s(x)\Big| > \lambda\Big\}\Big| \leq \frac{C}{\lambda}\|\varphi\|_1. \tag{2.3}$$

Proof. We will be able to use (2.2) to our advantage here, since each $b_{s,k}$ has mean 0 when averaged over dyadic intervals of size $2^{S(n-1)}$, and the Fourier bounds on μ_n are strong enough to exploit this.

We consider the standard ℓ^1 averages

$$\sigma_n = 2^{-S(n-1)-n} \chi_{[1, 2^{S(n-1)+n}]}, \tag{2.4}$$

and decompose $\mu_n = \mu_n * \sigma_n + \mu_n * (\delta_0 - \sigma_n)$. Accordingly, the set on the left of (2.3) is contained in the union of the sets

$$E_1 := \Big\{x : \sup_n \Big|(\mu_n * \sigma_n) * \sum_{s<S(n-1)} b_s(x)\Big| > \frac{\lambda}{2}\Big\},$$
$$E_2 := \Big\{x : \sup_n \Big|(\mu_n - \mu_n * \sigma_n) * \sum_{s<S(n-1)} b_s(x)\Big| > \frac{\lambda}{2}\Big\}.$$

Observe that for any $t > s$,

$$|\chi_{[1,2^t]} * b_{s,k}(x)| \le \begin{cases} 0, & x \notin Q_{s,k} + [0, 2^t], \\ 0, & x \in Q_{s,k} + [2^s, 2^t - 2^s], \\ \|b_{s,k}\|_1 & \text{otherwise,} \end{cases}$$

since each $b_{s,k}$ has mean 0 and is supported on $Q_{s,k}$. Therefore $\|\sigma_n * b_{s,k}\|_1 \le 2^{-S(n-1)-n+s+1}\|b_{s,k}\|_1$, which implies

$$\begin{aligned} |E_1| &\le 2\lambda^{-1} \sum_n \Big\|\mu_n * \sigma_n * \sum_{s<S(n-1)} b_s\Big\|_1 \le 2\lambda^{-1} \sum_n \|\mu_n\|_1 \sum_{s<S(n-1)} \|\sigma_n * b_s\|_1 \\ &\le 2\lambda^{-1} \sum_n \sum_{s<S(n-1)} 2^{-S(n-1)-n+s+1}\|b_s\|_1 \\ &\le 2\lambda^{-1} \sum_n \sum_s 2^{-n+1}\|b_s\|_1 \le \frac{C}{\lambda}\|\varphi\|_1. \end{aligned}$$

(This is a standard Calderón–Zygmund argument so far.)

Now for the other sum, we can write

$$1 - \hat{\sigma}_n(\gamma) = (1 - e(\gamma)) \sum_{j=0}^{2^{S(n-1)+n}-1} (1 - j2^{-S(n-1)-n})e(j\gamma),$$

and use (2.2) to bound

$$\|\hat{\mu}_n(1 - \hat{\sigma}_n)\|_\infty \le 2^{S(n-1)+n} \sup_\gamma |(1 - e(\gamma))\hat{\mu}_n(\gamma)| \le 2^{-S(n-1)-n}.$$

Here, in a variant of the technique from [4], we will use the extremely strong ℓ^2 estimate we get from this Fourier bound to obtain a weak ℓ^1 estimate. Starting with Chebyshev's inequality, we calculate

$$\begin{aligned} |E_2| &\le 4\lambda^{-2} \Big\|\sup_n \Big|(\mu_n - \mu_n * \sigma_n) * \sum_{s<S(n-1)} b_s\Big|\Big\|_2^2 \\ &\le 4\lambda^{-2} \sum_n \Big\|(\mu_n - \mu_n * \sigma_n) * \sum_{s<S(n-1)} b_s\Big\|_2^2 \\ &= 4\lambda^{-2} \sum_n \Big\|\hat{\mu}_n(1 - \hat{\sigma}_n) \sum_{s<S(n-1)} \hat{b}_s\Big\|_2^2 \end{aligned}$$

$$\begin{aligned}
&\le 4\lambda^{-2}\sum_n \|\hat{\mu}_n(1-\hat{\sigma}_n)\|_\infty^2 \Big\|\sum_{s<S(n-1)} \hat{b}_s\Big\|_2^2\\
&\le 4\lambda^{-2}\sum_n 2^{-2S(n-1)-2n}\Big\|\sum_{s<S(n-1)} b_s\Big\|_2^2\\
&= 4\lambda^{-2}\sum_n 2^{-2S(n-1)-2n}\sum_{s<S(n-1)} \|b_s\|_2^2 =: (*)
\end{aligned}$$

using the orthogonality of b_s for different s for the last step. Now since $\|b_s\|_\infty \le \|b_s\|_1 \le \lambda 2^s$, we get

$$\begin{aligned}
(*) &\le 4\lambda^{-2}\sum_n 2^{-2S(n-1)-2n}\sum_{s<S(n-1)} \lambda 2^s\|b_s\|_1\\
&\le 4\lambda^{-1}\sum_n 2^{-S(n-1)-2n}\sum_s \|b_s\|_1 \le \frac{C}{\lambda}\|\varphi\|_1.
\end{aligned}$$

This completes the proof. ∎

REMARK 2.2. This proof generalizes straightforwardly to measure-preserving $\mathbb{Z}^d$-actions, and indeed, to actions by finitely generated abelian groups (this requires defining the Calderón–Zygmund decomposition on such a group, using for instance the dyadic cubes from [5]). Note that the proof of Theorem 1.4 for $p = 2$ generalizes to this case, and thus we will have a.e. convergence of these ergodic averages for all $f \in L^1(X)$; see Theorem 2.4 in [1].

3. Threshold result: averages along $n^2 + \lfloor\rho(n)\rfloor$. In this section, we will prove Theorem 1.8.

We begin with the first claim, that if $\rho'(x) \gg x^{-1}$, then the $\{\mu_N\}$ have asymptotically trivial transforms. In this section, we will use the classical result of Weyl [14] on trigonometric sums, and we will repeatedly refer to its exposition in Section II.2 of Rosenblatt and Wierdl [11] rather than replicate it in its entirety here ([1]).

Let $\beta \in \mathbb{T}$. By Dirichlet's theorem on rational approximations, there exists a rational number p/q in lowest terms, with $q \le N^{4/3}$, such that

$$|\beta - p/q| \le q^{-1}N^{-4/3}. \tag{3.1}$$

We first write

$$\hat{\mu}_N(\beta) = \frac{1}{N}\sum_{j=0}^{\lfloor\rho(N)\rfloor} e(j\beta)\sum_{k\in I_j} e(k^2\beta) + O\bigg(\frac{L_{\lfloor\rho(N)\rfloor}}{N}\bigg),$$

([1]) Many theorems and equation numbers in this paper will refer to [11], rather than to Section 2 above. Fortunately, none of the theorems or equation numbers will coincide between our Section 2 and their Section II.2.

where $I_j := \{x \in \mathbb{R}^+ : \lfloor \rho(k) \rfloor = j\}$, and denote $L_j := |I_j|$. Note that by the hypotheses on ρ, we have L_j increasing and $j^{\epsilon+2/3} \lesssim L_{\lfloor \rho(j) \rfloor} \ll j$ as $j \to \infty$. We adapt from [11] the notation

$$\hat{A}(p/q) := \frac{1}{q} \sum_{n=0}^{q-1} e(n^2 p/q)$$

for $p/q \in \mathbb{Q}$, as well as the functions

$$V_j(\alpha) := \sum_{l:\, \sqrt{l} \in I_j} \frac{1}{2\sqrt{l}} e(l\alpha)$$

for any $\alpha \in \mathbb{R}$.

As is standard in the Hardy–Littlewood circle method, we will make use of different estimates depending on the size of the denominator q relative to N.

Lemma 3.1. *There exists a constant $C < \infty$, depending only on ρ and ϵ, such that for $j > \rho(N^{1-\epsilon})$, $q \le N^{2/3}$, and $|\beta - p/q| \le q^{-1} N^{-4/3}$, we have*

$$\Big| \sum_{k \in I_j} e(k^2 \beta) - \hat{A}(p/q) V_j(\beta - p/q) \Big| \le C N^{-\epsilon/6} L_j. \tag{3.2}$$

Similarly, if $j > \rho(N^{1-\epsilon})$, $N^{2/3} < q \le N^{4/3}$, and $|\beta - p/q| \le q^{-1} N^{-4/3}$, then

$$\Big| \sum_{k \in I_j} e(k^2 \beta) \Big| \le C N^{-\epsilon/7} L_j. \tag{3.3}$$

Proof. (3.2) is the equivalent of (2.25) of [11], replacing the exponential sums beginning at 0 with the sums along the interval $\{l : \sqrt{l} \in I_j\}$. As in (2.27),

$$\Big| \sum_{k \in I_j} e(k^2 p/q) - \hat{A}(p/q) V_j(0) \Big| \le Cq$$

clearly holds with a universal constant. Thus we may apply Lemma 2.13 (note that we are summing in l, over $\le 2NL_j$ terms) to find

$$\Big| \sum_{k \in I_j} e(k^2 \beta) - \hat{A}(p/q) V_j(\beta - p/q) \Big| \le Cq(NL_j |\beta - p/q| + 1) \le C N^{-\epsilon/6} L_j,$$

using for the Cq term the assumption $q \le N^{2/3}$ and the fact that for $j > \rho(N^{1-\epsilon})$, we have

$$L_j \gtrsim N^{(1-\epsilon)(2/3+\epsilon)} \gtrsim N^{2/3+\epsilon/6}.$$

Similarly, (3.3) is the analogue of (2.28), and the required estimate

$$\Big| \sum_{k \in I_j} e(k^2 p/q) \Big| \le C \Big(\frac{L_j}{\sqrt{q}} + \sqrt{q \log L_j} \Big),$$

like (2.29), relies only on the fact that the squares are summed along an interval. We therefore apply Lemma 2.13 to obtain

$$\Big|\sum_{k\in I_j} e(k^2\beta)\Big| \le C\bigg(\frac{L_j}{\sqrt{q}} + \sqrt{q\log L_j}\bigg)(NL_j|\beta - p/q| + 1),$$

which for $q > N^{2/3}$ is indeed bounded by $CN^{-\epsilon/7}L_j$. ∎

Thus we have a satisfactory bound for $q > N^{2/3}$, while for $q \le N^{2/3}$ and $|\beta - p/q| \le q^{-1}N^{-4/3}$, we have

$$\hat{\mu}_N(\beta) = \frac{1}{N}\sum_{j=0}^{\lfloor\rho(N)\rfloor} e(j\beta)\hat{\Lambda}(p/q)V_j(\beta - p/q) + O\bigg(\frac{L_j}{N} + N^{-\epsilon/6}\bigg).$$

Now

$$\frac{1}{N}\sum_{j=0}^{\lfloor\rho(N)\rfloor} e(j\beta)V_j(\beta - p/q)$$

is a Cesàro mean of the averages

$$\frac{1}{N^2}\sum_{j=0}^{\lfloor\rho(N)\rfloor} e(j\beta)\sum_{l:\sqrt{l}\in I_j} e(l(\beta - p/q)), \tag{3.4}$$

and so it suffices for Theorem 1.8 to show that these averages decay to 0 (away from $\beta = 0$) at a rate depending only on ρ. Let $\alpha = \beta - p/q$. For $\alpha \ne 0$,

$$\frac{1}{N^2}\sum_{j=0}^{\lfloor\rho(N)\rfloor} e(j\beta)\sum_{l:\sqrt{l}\in I_j} e(l\alpha) = \frac{1}{N^2}\sum_{j=0}^{\lfloor\rho(N)\rfloor} e(j\beta)\frac{e(\lceil\phi(j+1)\rceil\alpha) - e(\lceil\phi(j)\rceil\alpha)}{e(\alpha) - 1}$$

where $\phi(j) := (\rho^{-1}(j))^2$. We use the sum version of integration by parts:

$$\sum_{j=0}^{m}(a_{j+1} - a_j)b_j = a_{m+1}b_m - a_1b_0 + \sum_{j=0}^{m} a_j(b_{j-1} - b_j),$$

with

$$a_j := N^{-2}\sum_{i=0}^{j-1} e(i\beta) \quad\text{and}\quad b_j := \frac{e(\lceil\phi(j+1)\rceil\alpha) - e(\lceil\phi(j)\rceil\alpha)}{e(\alpha) - 1}.$$

We evaluate the end terms first, noting that

$$\begin{aligned}|a_{j+1}b_j| &\le \frac{2}{N^2|\beta|}\bigg|\frac{e(\lceil\phi(j+1)\rceil\alpha) - e(\lceil\phi(j)\rceil\alpha)}{e(\alpha) - 1}\bigg| \\ &\le \frac{C(\phi(j+1) - \phi(j))}{N^2|\beta|} = O\bigg(\frac{L_j}{N|\beta|}\bigg).\end{aligned}$$

Now the main sum is

$$\begin{aligned}
&\Big|\sum_{j=0}^{\lfloor\rho(N)\rfloor} a_j(b_{j-1}-b_j)\Big| \\
&\quad= \left|\frac{1}{N^2}\sum_{j=0}^{\lfloor\rho(N)\rfloor}\frac{1-e(j\beta)}{1-e(\beta)}\,\frac{e(\lceil\phi(j+1)\rceil\alpha)-2e(\lceil\phi(j)\rceil\alpha)+e(\lceil\phi(j-1)\rceil\alpha)}{1-e(\alpha)}\right| \\
&\quad\le \frac{2}{N^2|\beta|\,|\alpha|}\Big|\sum_{j=0}^{\lfloor\rho(N)\rfloor} e(\lceil\phi(j+1)\rceil\alpha)-2e(\lceil\phi(j)\rceil\alpha)+e(\lceil\phi(j-1)\rceil\alpha)\Big| \\
&\quad\le \frac{C}{N^2|\beta|\,|\alpha|}\sum_{j=0}^{\lfloor\rho(N)\rfloor}\phi''(j+1)|\alpha| \le \frac{C\phi'(N)}{N^2|\beta|} = O\Big(\frac{L_{\lfloor\rho(N)\rfloor}}{N|\beta|}\Big),
\end{aligned}$$

using the monotonicity of ρ and its derivatives (and thus the monotonicity of ϕ'') to justify the second inequality.

If $\alpha=0$, then by the same methods,

$$\begin{aligned}
\frac{1}{N^2}\sum_{j=0}^{\lfloor\rho(N)\rfloor} e(j\beta)|\{l:\sqrt{l}\in I_j\}| &= \frac{1}{N^2}\sum_{j=0}^{\lfloor\rho(N)\rfloor} e(j\beta)(\lceil\phi(j+1)\rceil-\lceil\phi(j)\rceil) \\
&= O\Big(\frac{L_{\lfloor\rho(N)\rfloor}}{N|\beta|}\Big).
\end{aligned}$$

Therefore we have proved that

$$|\hat{\mu}_N(\beta)| \lesssim N^{-\epsilon/7} + \frac{L_{\lfloor\rho(N)\rfloor}}{N|\beta|}, \tag{3.5}$$

which clearly establishes that the sequence $\{\mu_N\}$ has asymptotically trivial transforms.

Now we turn to the second claim of Theorem 1.8, that if $\rho'(x)\le Cx^{-1}$ (thus $\rho(x)\le C\log x$), then no subsequence of the averages along $k^2+\lfloor\rho(k)\rfloor$ satisfies a weak-type $(1,1)$ maximal inequality, and thus by the Conze principle [6], no subsequence of these averages can converge a.e. for all L^1 functions in an ergodic dynamical system. This follows from the argument of [9], in which it was proved that the same is true of the sequence k^2 among others. (That paper is an extension in several directions of the paper of Buczolich and Mauldin [3], which proved that the full sequence of averages along k^2 does not satisfy a weak maximal inequality.)

Given such a ρ and any subsequence $\{N_k\}$ of the averages, we choose a further subsequence $\{k_i\}$ such that modulo any squarefree odd Q, $\lfloor\rho(\frac{1}{2}N_{k_i})\rfloor$ has a limit r_Q as $i\to\infty$. (This is done by a diagonal argument, since we only need to ensure that this happens modulo the product of the first M

primes, for each M.) Then if we restrict to this subsequence, we see that the quadratic residues translated by r_q serve as the Λ_q in Theorem 3.1 of [9], that (3.1)–(3.4) hold for the same reasons as for the original quadratic residues, and that (3.5) holds along our subsequence. That is, for any squarefree and odd Q with sufficiently large factors and any non-trivial quadratic residue a modulo Q,

$$\begin{aligned}&\liminf_{k\to\infty}\frac{1}{N_k}|\{1\le j\le N_k : j^2+\lfloor\rho(j)\rfloor\equiv a+r_Q \bmod Q\}|\\&\quad\ge\liminf_{k\to\infty}\frac{1}{N_k}|\{1\le j\le N_k : j^2\equiv a \bmod Q, \lfloor\rho(j)\rfloor=\lfloor\rho(N_k/2)\rfloor\}|\\&\quad\ge\liminf_{k\to\infty}|\Lambda_Q|^{-1}\frac{1}{2N_k}|\{1\le j\le N_k : \lfloor\rho(j)\rfloor=\lfloor\rho(N_k/2)\rfloor\}|-\frac{1}{Q}\ge\frac{1}{3C|\Lambda_Q|},\end{aligned}$$

using the fact that $\rho'(x)\le 4C/N$ for all $x\ge N/4$, so that $|\{1\le j\le N_k : \lfloor\rho(j)\rfloor=\lfloor\rho(N_k/2)\rfloor\}|\ge\min\{N/4C, N/4\}$, and the fact that $|\Lambda_Q|\ll Q$ for Q large. Note that Theorem 4.1 in [9] implies this variant of Theorem 3.1, since (4.7) is the only use of (3.5) in that paper. Therefore any subsequence of the μ_N has a further subsequence for which the weak L^1 maximal inequality fails, which implies our desired result.

Acknowledgements. The author thanks J. Rosenblatt for introducing him to these interesting problems, for pointing out the article [8], and for many suggestions; and his dissertation advisor, M. Christ, for several major ideas, including treating the intermediate terms in Theorem 1.5 separately after the fashion of [12].

This research was partly supported by the NSF (Grant DMS-0401260.).

REFERENCES

[1] A. Bellow, R. L. Jones, and J. Rosenblatt, *Almost everywhere convergence of weighted averages*, Math. Ann. 293 (1992), 399–426.

[2] M. Boshernitzan, G. Kolesnik, A. Quas, and M. Wierdl, *Ergodic averaging sequences*, J. Anal. Math. 95 (2005), 63–103.

[3] Z. Buczolich and R. D. Mauldin, *Divergent square averages*, Ann. of Math. 171 (2010), 1479–1530.

[4] M. Christ, *Weak type* $(1,1)$ *bounds for rough operators*, ibid. 128 (1988), 19–42.

[5] —, *A* $T(b)$ *theorem with remarks on analytic capacity and the Cauchy integral*, Colloq. Math. 60/61 (1990), 601–628.

[6] J.-P. Conze, *Convergence des moyennes ergodiques pour des sous-suites*, in: Contributions au calcul des probabilités, Bull. Soc. Math. France Mém. 35 (1973), 7–15.

[7] C. Fefferman, *Inequalities for strongly singular convolution operators*, Acta Math. 124 (1970), 9–36.

[8] S. Kostyukovsky and A. Olevskii, *Compactness of families of convolution operators with respect to convergence almost everywhere*, Real Anal. Exchange 30 (2004/05), 755–765.

[9] P. LaVictoire, *Universally L^1-bad arithmetic sequences*, J. Anal. Math. 113 (2011), 241–263.
[10] J. M. Rosenblatt, *Convergence of sequences of convolution operators*, New Zealand J. Math. 38 (2008), 137–147.
[11] J. M. Rosenblatt and M. Wierdl, *Pointwise ergodic theorems via harmonic analysis*, in: Ergodic Theory and Its Connections with Harmonic Analysis (Alexandria, 1993), London Math. Soc. Lecture Note Ser. 205, Cambridge Univ. Press, Cambridge, 1995, 3–151.
[12] A. Seeger, T. Tao, and J. Wright, *Singular maximal functions and Radon transforms near L^1*, Amer. J. Math. 126 (2004), 607–647.
[13] R. Urban and J. Zienkiewicz, *Weak type* $(1, 1)$ *estimates for a class of discrete rough maximal functions*, Math. Res. Lett. 14 (2007), 227–237.
[14] H. Weyl, *Über die Gleichverteilung von Zahlen mod. Eins*, Math. Ann. 77 (1916), 313–352.

Patrick LaVictoire
Department of Mathematics
University of Wisconsin at Madison
Madison, WI 53706-1388, U.S.A.
E-mail: patlavic@math.wisc.edu

Received 16 September 2010 (5428)