ILLINOIS JOURNAL OF MATHEMATICS
Volume 37, Number 4, Winter 1993

ALMOST EVERYWHERE CONVERGENCE OF CONVOLUTION POWERS IN $L^1(X)$

KARIN REINHOLD-LARSSON

Introduction

This paper is concerned with the behavior of weighted averages induced by a probability measure on the integers. Let (X, β, m) be a probability space and τ: $X \mapsto X$ an invertible measure preserving point transformation. A probability measure μ on $\mathbf{Z}$, the integers, gives rise to the weighted average

$$\mu f(x) = \sum_{k=-\infty}^{\infty} \mu(k) f(\tau^k x).$$

The powers of the operator μf are defined by the convolution powers of the measure μ,

$$\mu^n f(x) = \sum_{k=-\infty}^{+\infty} \mu^n(k) f(\tau^k x)$$

where, on the right hand side, $\mu^n(k)$ denotes the nth convolution power of μ evaluated at k. Note that since

$$\left(\int |\mu^n f(x)|^p \, dm(x)\right)^{1/p} \leq \sum_{k \in \mathbf{Z}} \mu(k) \left(\int |f(\tau^k x)|^p \, dm(x)\right)^{1/p} = \|f\|_p,$$

these operators are well defined a.e. and are positive contractions in all $L^p(X)$, $1 \leq p \leq \infty$. Bellow-Jones-Rosenblatt [2], [3], [5] studied these types of averaging operators as well as more general types of weighted averages. They proved these operators converge in norm whenever the support of μ is not contained in a coset of a proper subgroup of $\mathbf{Z}$. In addition they proved [3] that if the measure is centered and has finite second moment then there is convergence almost everywhere in $L^p(X)$ for all $p > 1$. Their method is based on Fourier techniques that could not be extended to L^1. V. I. Oseledec [14] proved convergence almost everywhere in L^1 for symmetric

Received December 30, 1991.
1991 Mathematics Subject Classification. Primary 28 Dxx; Secondary 47A35.

measures without any moment condition. His proof is based on Doob's Martingale Convergence Theorem. The main result of this paper shows a.e. convergence in $L^1(X)$, via estimations of the probability distribution of $\mu^n(k), k \in \mathbf{Z}$.

This paper is divided into two parts. The first one deals with a.e. convergence of $\mu^n f$ in $L^1(X)$ for measures that are centered and satisfy some moment condition, so that their support has an appropriate distribution. The second one analyzes the a.e. convergence of $\mu^n f$ along subsequences for measures μ which are not centered. For such measures, convergence of the whole sequence does not hold but a moment condition is still required for a subsequence result.

The results of this paper form part of the author's thesis. She would like to express her indebtedness to her advisor, Professor Joseph Rosenblatt, who suggested the problem and whose support and guidance gave constant encouragement.

1. Convergence for centered measures

1.1. Probability measures on Z

Initially, we will focus on the properties of the measure μ, beginning with some definitions and well-known facts.

DEFINITION 1.1.2. Let μ be a probability measure on $\mathbf{Z}$. We say that μ is *adapted* if its support generates $\mathbf{Z}$, and that μ is *strictly aperiodic* if its support is not contained in a coset of a proper subgroup of $\mathbf{Z}$. For $\alpha > 0$, the *α-moment* of μ is defined as $\sum_{k=-\infty}^{\infty} |k|^\alpha \mu(k)$, and is denoted by $m_\alpha(\mu)$. The *expected value* of μ is $\sum_{k=-\infty}^{\infty} k\mu(k)$, and is denoted by $E(\mu)$. The measure μ is called *centered* if it has expected value zero.

The following are useful characterizations of strictly aperiodic probabilities.

PROPOSITION 1.1.2 (Foguel [8]). *Let μ be a probability measure on $\mathbf{Z}$.*

(i) *If μ is adapted, then μ is strictly aperiodic if and only if* $\lim_{n\to\infty} \|\mu^{n+1} - \mu^n\|_{l^1(\mathbf{Z})} = 0$.

(ii) *μ is strictly aperiodic if and only if $|\hat{\mu}(\lambda)| < 1$ for all $\lambda \in \mathbf{C}$ with* $|\lambda| = 1$, $\lambda \neq 1$.

The strict aperiodicity of the measure is needed to prove convergence a.e. for all functions on $L^1(X)$. However, for convergence in norm, one only needs the measure to be adapted.

PROPOSITION 1.1.3 (Bellow, Jones and Rosenblatt [3]). *Let $1 \le p < \infty$. Then*

$$\{f \in L^p(X): \mu f = f\} + \mathrm{cl}\{f - \mu f: f \in L^p(X)\}$$

is a dense subspace of $L^p(X)$. Also, if μ is adapted and τ is ergodic, $\mathrm{cl}\{f - \mu f: f \in L^p(X)\}$ is the subspace of mean zero functions in $L^p(X)$.

From these two propositions it follows the convergence in norm.

COROLLARY 1.1.4. *If μ is strictly aperiodic, then for every $f \in L^1(X)$, $\lim_{n\to\infty}\|\mu^n f - P_I f\|_1 = 0$, where $P_I f \in L^1(X)$ is the projection of f onto the subspace of μ-invariant functions. In particular, if μ is strictly aperiodic and τ is ergodic, $\lim_{n\to\infty}\|\mu^n f - \int f\,dm\|_1 = 0$.*

Also, from Proposition 1.1.3, pointwise convergence on a dense subspace follows because

$$\{f - \mu f: f \in L^\infty(X)\} + \{f \in L^1(X): \mu f = f\}$$

is a dense subspace of $L^1(X)$.

PROPOSITION 1.1.5. *If μ is strictly aperiodic, then for f in a dense subspace of $L^1(X)$, $\lim_{n\to\infty}\mu^n f(x)$ exist for almost every x.*

Applying the Banach Principle, and since there is a.e. convergence on a dense subspace, it remains to prove that the operator $\sup_{n\in\mathbf{N}}|\mu^n f(x)|$ is of weak type (1, 1); that is, it suffices to show that there is a constant $C > 0$ such that

$$m\left\{x \in X: \sup_{n\in\mathbf{N}} |\mu^n f(x)| > \lambda\right\} \le C\frac{\|f\|_1}{\lambda} \tag{1.a}$$

for all $f \in L^1(X)$. This weak (1, 1) maximal inequality is proved by comparing the distribution of μ^n with that of the nth-convolution power of the Gaussian distribution.

THEOREM 1.1.6. *Let μ be a strictly aperiodic probability measure on the integers with $m_{2+\delta}(\mu) < \infty$ for some $0 < \delta \le 1$. Then,*

$$\sup_{k\in\mathbf{Z}}\left|\mu^n(k) - \frac{1}{\sigma\sqrt{2\pi n}}\exp\left(-\frac{(k-an)^2}{2\sigma^2 n}\right)\right| \le \frac{C}{n^{(1+\delta)/2}} \tag{1.b}$$

where a denotes the expected value of μ, $a = E(\mu)$.

This limit theorem is a classical result in the theory of infinitely divisible distributions. A complete exposition of limit theorems for probabilities in the domain of attraction of infinite divisible distributions can be found in Ibragimov-Linnik [11]. The proof of this theorem is omitted since its proof is essentially in Ibragimov-Linnik [11], Theorem 4.2.1, page 121, Theorem 4.5.3, page 138. Their argument can be extended to measures with a finite moment bigger than 2. The strict aperiodicity of the measure μ simplifies some of the calculations.

Notes. (1) If the measure μ had a moment higher than 3, the rate of decay of the difference between the convolution powers of μ and the convolution powers of the normal distribution would not be faster than $1/n$. This rate could be improved in case the third derivative of the characteristic function of μ vanishes at 0. Symmetric measures with a finite third moment have a vanishing third derivative but it is desirable to avoid, if possible, imposing such a restrictive condition on μ.

(2) For $\delta = 0$, the above theorem does not give an uniform estimate on $\mu^n(k)$. However, it is interesting to see that its rate of decay is of order $n^{-1/2}$. Chung and Erdös [7] have the following surprising result.

LEMMA 1.1.7. *Let μ be a strictly aperiodic probability measure on $\mathbf{Z}$ such that $\{k:\mu(k) > 0\}$ do not have all the same sign. If $m_1(\mu) < \infty$,*

$$\sup_{k \in \mathbf{Z}} \mu^n(k) \leq \frac{C}{\sqrt{n}}$$

where C does not depend on n.

However, technicalities force one to ask moments higher than 2.

With the limit theorem 1.1.6, one can prove the maximal estimate (1.a).

THEOREM 1.1.8. *If μ has finite support and $E(\mu) = 0$ then $\sup_{n \in \mathbf{N}} |\mu^n f(x)|$ is a weak $(1, 1)$ operator.*

Proof. Suppose $\operatorname{supp}(\mu) \subseteq [-N, N]$, for some positive integer N. Let ϕ_n be a discrete version of the Gaussian distribution, i.e.

$$\phi_n(k) = \begin{cases} \dfrac{1}{\sigma\sqrt{2\pi n}} \exp\left(-\dfrac{k^2}{2\sigma^2 n}\right) & \text{for } k \in [-nN, nN] \\ 0 & \text{otherwise.} \end{cases}$$

From Theorem 1.1.6, one has

$$|\mu^n(k) - \phi_n(k)| \le \frac{C}{n} \quad \text{for all } k \in [-nN, nN].$$

(Theorem 1.1.6 was used with $\delta = 1$, since μ has finite support and certainly, finite third moment.) In other words, denoting with

$$A_n = \frac{1}{2n+1} \sum_{j=-n}^{n} \delta_{\{j\}}$$

the measure corresponding to the usual averages,

$$|\mu^n(k) - \phi_n(k)| \le \frac{C}{n} \sum_{j=-nN}^{nN} \delta_{\{j\}}(k) = C\frac{2nN+1}{n} A_{nN}(k).$$

Therefore, since A_n satisfies a maximal inequality, so does $|\mu^n - \phi_n|$.

The operator $\sup_{n \in \mathbf{N}} |\phi_n f|$ is also weak (1, 1) because the ϕ_n's are radially decreasing. In fact, each of them can be rewritten as an almost convex combination of usual averages:

$$\phi_n = \sum_{k=0}^{nN} a_k^n A_k \quad \text{where } \sum_{k=0}^{nN} a_k^n \le 2 \text{ and } a_k^n > 0 \text{ for } 0 \le k \le nN.$$

Hence, one can show that

$$\sup_{n \in \mathbf{N}} |\phi_n f(x)| \le 2 \sup_{n \in \mathbf{N}} |A_n f(x)|.$$

The details are left to the reader.

Writing $\mu^n = (\mu^n - \phi_n) + \phi_n$, it follows that $\sup_{n \in \mathbf{N}} |\mu^n f|$ is also a weak (1, 1) operator. □

Note. The condition $E(\mu) = 0$ was used in that the ϕ_n's and the A_n's can be centered at the origin. The maximal inequalities for A_n and ϕ_n would not hold otherwise.

Theorem 1.1.8 can be extended to measures which do not have finite support. However, in such cases, a moment condition is required. This condition arises from the estimation of the tail distribution of the powers of μ, which is translated to estimating large deviation probabilities. There has been a great deal of study of such problem. For our purpose, we use the following theorem of Baum and Katz which gives the lowest moment condition for μ.

THEOREM 1.1.9 (Baum and Katz [1]). *Let $\{X_i\}$ be a family of i.i.d. random variables with $E(X_1) = 0$. If $t > 1$, $r > 1$ and $1/2 < r/t \leq 1$, the following are equivalent:*

(i) $E(|X_1|^t) < \infty$ (*finite* tth *moment*),

(ii) $\sum_{n=1}^{\infty} n^{r-2} P(|S_n| > n^{r/t}\epsilon) < \infty$ *for all* $\epsilon > 0$.

PROPOSITION 1.1.10. *If $E(\mu) = 0$ and $m_{2+\delta}(\mu) < \infty$ for some $\delta \geq (\sqrt{17} - 3)/2$, then $\sup_{n \in \mathbf{N}} |\mu^n f|$ is a weak $(1,1)$ operator.*

Proof. Let $\alpha = (1 + \delta)/2$ and split μ^n into two pieces, $\mu^n = \nu_n + \omega_n$ where $\nu_n = \mu^n|_{[-[n^\alpha],[n^\alpha]]}$ and $\omega_n = \mu^n - \nu_n$. The center piece, ν_n, is handled in the same way as μ^n was in Theorem 1.1.8, that is,

$$|\nu_n(k) - \phi_n(k)| \leq \frac{C}{[n^\alpha]}$$

for all $|k| \leq [n^\alpha]$, where ϕ_n is the discrete version of the Gaussian distribution, now with support in $[-[n^\alpha], [n^\alpha]]$. Arguing as in Theorem 1.1.8,

$$\begin{aligned}\sup_{n \in \mathbf{N}} |\nu_n f(x)| &\leq \sup_{n \in \mathbf{N}} |\nu_n f(x) - \phi_n f(x)| + \sup_{n \in \mathbf{N}} |\phi_n f(x)| \\ &\leq C \sup_{n \in \mathbf{N}} A_{[n^\alpha]}|f|(x) + \sup_{n \in \mathbf{N}} |\phi_n f(x)| \\ &\leq C' \sup_{n \in \mathbf{N}} A_n |f|(x),\end{aligned}$$

proving that $\sup_{n \in \mathbf{N}} |\nu_n f|$ satisfies a weak $(1,1)$ inequality.

The tails, ω_n, are controlled by estimating

$$\sup_{n \in \mathbf{N}} |\omega_n f(x)| \leq \sum_{n=0}^{\infty} |\omega_n f(x)|,$$

which gives

$$m\left\{x \in X: \sup_{n \in \mathbf{N}} |\omega_n f(x)| > \lambda\right\} \leq \frac{1}{\lambda} \|f\|_1 \sum_{n=0}^{\infty} \|\omega_n\|_1.$$

This is the point where the moment condition for μ is needed. Thinking of μ^n as the distribution of the sum of n i.i.d. random variables, say $\{X_i\}$, each with distribution μ, the l_1-norms of the tails are

$$\begin{aligned}\|\omega_n\|_1 &= \sum_{|k| > [n^\alpha]} \mu^n(k) = P(|X_1 + X_2 + \cdots + X_n| > [n^\alpha]) \\ &= P(|S_n| > [n^\alpha]).\end{aligned}$$

Therefore, the problem is transformed to estimate large deviation probabilities. In view of Baum and Katz's theorem, take $r = 2$ and $r/t = \alpha = (1 + \delta)/2$. Therefore, $t = 4/(1 + \delta)$. Since $m_{2+\delta}(\mu) < \infty$, then $t \leq 2 + \delta$, therefore δ must satisfy $4/(1 + \delta) \leq 2 + \delta$. Solving this equation we have that whenever $\delta \geq (\sqrt{17} - 3)/2$, $\sum_{n=0}^{\infty} \|\omega_n\|_1 < \infty$, and hence, $\sup_{n \in \mathbf{N}} |\omega_n f|$ is integrable. □

This concludes the proof of the following theorem:

THEOREM 1.1.11. *Let μ be a strictly aperiodic probability measure on $\mathbf{Z}$ with $E(\mu) = 0$ and $m_{2+\delta}(\mu) < \infty$ for some $\delta \geq (\sqrt{17} - 3)/2$. Then, for any $f \in L^1(X)$,*

$$\lim_{n \to \infty} \mu^n f(x) = P_I f(x) \quad \text{for a.e. } x.$$

For measures with a finite moment bigger than two there is, nevertheless, a subsequence result. Indeed, one can show that there is a sequence of polynomial growth, under which there is convergence almost everywhere. Notice that, for the usual averages, it suffices to prove convergence a.e. along an exponential sequence. The next proposition shows convergence along a considerably "thicker" sequence for the convolution operators, but nevertheless, this does not give convergence of the full sequence.

PROPOSITION 1.1.12. *Let μ be a strictly aperiodic probability measure on $\mathbf{Z}$ with $m_{2+\delta}(\mu) < \infty$ for some $\delta > 0$ and $E(\mu) = 0$. Then, there exist two positive integers $k = k(\delta)$ and $N = N(k)$, such that $\lim_{n \to \infty} \mu^{n^k} f(x)$ exists a.e. for all $f \in L^1(X)$ and $\lim_{n \to \infty} \mu^n f(x)$ exists a.e. for all $f \in L \log^{N-1} L$.*

Proof. Let ω_n denote the tail of μ^n as in the previous proposition. It suffices to prove the existence of an integer $k = k(\delta)$ such that

$$\sum_{n=1}^{\infty} \|\omega_{n^k}\|_1 < \infty.$$

Since $E(|S_n|^2) \leq nE(|X_1|^2)$, then

$$\begin{aligned} \|\omega_{n^k}\|_1 = P\left(|S_{n^k}| > (n^k)^{(1+\delta)/2}\right) &\leq \frac{1}{n^{k(1+\delta)}} E\left(|S_{n^k}|^2\right) \\ &\leq K \frac{n^k}{n^{k(1+\delta)}} = K \frac{1}{n^{k\delta}}. \end{aligned}$$

Therefore, if $k > 1/\delta$, then $\sum_{n=1}^{\infty} \|\omega_{n^k}\|_1 < \infty$.

For the second statement, observe that for fixed k, there exists $N = N(k)$, such that any positive integer can be written as the sum of at most N powers of k, i.e., $n = n_1^k + n_2^k + \cdots + n_r^k$ with $r \leq N, n, n_1, \cdots, n_r \in \mathbf{N}$. This is the Waring-Hilbert Theorem; see Hua [10], chapters 18-19. By replacing f with $|f|$ one can assume $f \geq 0$. Then,

$$\sup_{n \in \mathbf{N}} \mu^n f(x) \leq \sup_{n_1 \geq 0} \sup_{n_2 \geq 0} \cdots \sup_{n_N \geq 0} \mu^{n_1^k} \mu^{n_2^k} \cdots \mu^{n_N^k} f(x).$$

But for any M, if $f \in L \log^M L$ then $\sup_{n \in \mathbf{N}} A_n f \in L \log^{M-1} L$ where $A_n f$ denotes the usual average as in Theorem 1.1.8 (see Krengel [12, page 54]). Let ν_n be the central part of μ^n as in Proposition 1.1.10. Then $\sup_{n \in \mathbf{N}} \nu_{n^k} f$ is dominated by $\sup_{n \in \mathbf{N}} A_n f$ and $\sup_{n \in \mathbf{N}} \omega_{n^k} f$ is integrable. Therefore, if $f \in L \log^{N-1} L$, then

$$\sup_{n_2 \geq 0} \cdots \sup_{n_N \geq 0} \mu^{n_2^k} \cdots \mu^{n_N^k} f \in L^1$$

and

$$\int \sup_{n_2 \geq 0} \cdots \sup_{n_N \geq 0} \mu^{n_2^k} \ldots \mu^{n_N^k} f(x) dx \leq C \int f(x) \log^{N-1}(f(x))\, dx.$$

Consequently,

$$m\left\{x : \sup_{n_1 \geq 0} \sup_{n_2 \geq 0} \cdots \sup_{n_N \geq 0} \mu^{n_1^k} \mu^{n_2^k} \cdots \mu^{n_N^k} f(x) > \lambda\right\}$$

$$\leq C \int f(x) \log^{N-1}(f(x))\, dx. \quad \square$$

Remark. With the same techniques, one can prove that if μ is a strictly aperiodic probability measure on $\mathbf{Z}$ with $m_{2+\delta}(\mu) < \infty$ for some $\delta > 0$ and $E(\mu) = 0$, then $\lim_{n \to \infty} \mu^{2^n} f(x)$ exists a.e. for all $f \in L^1(X)$.

1.2. Extensions to $\mathbf{Z}^d$

The methods employed for measures on $\mathbf{Z}$ can be used to extend the results to measures on locally compact subgroups of $\mathbf{R}^d$.

DEFINITION 1.2.1. Let μ be a probability measure on an abelian, locally compact group G with Haar measure λ_G. We say that μ is *adapted* if its support generates G, and that μ is *strictly aperiodic* if its support is not contained in a coset of a proper closed subgroup of G. Also, μ is called *spread out* if there exists n such that μ^n and λ_G are not mutually singular.

Notice that if G is discrete, then μ is always spread-out. For subgroups of $\mathbf{R}^d$, the expected value and the moments of a measure can be defined as follows. Let $\|\mathbf{x}\|$ denote the Euclidean norm in $\mathbf{R}^d$.

DEFINITION 1.2.2. If G is a subgroup of $\mathbf{R}^d$ and $\alpha > 0$, the *α-moment* of μ is defined as $\int_G \|g\|^\alpha d\mu(g)$, and is denoted by $m_\alpha(\mu)$. If g_i is the ith coordinate of $g \in \mathbf{R}^d$, $g = (g_1, \cdots, g_d)$, define $a_i = \int_G g_i d\mu(g)$ to be the partial expected values of μ. Then $\mathbf{a} = (a_1, \cdots, a_d)$ is the *(vector-)expected value* of μ. We say that μ is *centered* if $\mathbf{a} = 0$.

PROPOSITION 1.2.3. *Let μ be a strictly aperiodic measure on $\mathbf{Z}^d$ with $\mathbf{a} = 0$. If d and μ satisfy one of the following three cases then* $\lim_{n\to\infty} \mu^n f(x)$ *exists a.e. for all $f \in L^1(X)$.*

(1) *$d = 1, 2$ and $m_{2+\delta}(\mu) < \infty$, for some $[\sqrt{(2+d)^2 + 8d} - (d+2)]/2 \leq \delta$.*

(2) *$d = 3$ and $m_3(\mu) < \infty$.*

(3) *$d > 3$, μ is symmetric and $m_{3+\delta}(\mu) < \infty$ for some δ in the range*

$$\left[\sqrt{(2+d)^2 + 8d} - (d+2)\right]/2 \leq \delta < 1.$$

The proof of this proposition follows the same line of argument as those employed in Section 1.1.

For measures on $\mathbf{R}^d$, the analogous proposition holds under the additional hypothesis that μ is spread out and there exists $n_0 \in \mathbf{N}$ for which $d\mu^{n_0}\mathbf{x} = l(\mathbf{x})d\mathbf{x}$ and $l \in L^p(\mathbf{R}^d) \cap L^2(\mathbf{R}^d)$.

1.3. On the maximal function

For measures μ on $\mathbf{Z}$ which satisfy the hypothesis of Theorem 1.1.11, the maximal function $\sup_{n\in\mathbf{N}} \mu^n|f|$ has the same properties as the maximal function corresponding to the usual averages, $\sup_{n\in\mathbf{N}} A_n|f|$. Bellow-Jones-Rosenblatt [3] proved that whenever μ is a strictly aperiodic centered probability measure on $\mathbf{Z}$ with $m_2(\mu) < \infty$, the maximal function $\sup_{n\in\mathbf{N}}|\mu^n f|$ is a strong type (p, p) operator for all $p > 1$. For $p = 1$, Theorem 1.1.6 and Proposition 1.1.10 yield the following results.

THEOREM 1.3.1. *Let μ be a strictly aperiodic, centered probability measure on $\mathbf{Z}$, with $m_{2+\delta}(\mu) < \infty$ for some $\delta > 0$; and let τ be an ergodic measure preserving invertible transformation. Then:*

(1) $\sup_{n\in\mathbf{N}} \mu^n|f| \in L^1(X) \Rightarrow f \in L \log L$;

(2) *if $f \in L \log L$ and $\delta > (\sqrt{17} - 3)/2$ then* $\sup_{n\in\mathbf{N}}|\mu^n f| \in L^1(X)$.

LEMMA 1.3.2. *Let μ be a strictly aperiodic probability measure on* $\mathbf{Z}$ *with* $E(\mu) = 0$ *and* $m_{2+\delta}(\mu) < \infty$ *for some* $\delta \geq 0$. *Then, for all* $\alpha \geq \min(0,(2 - \delta^2 - 3\delta)/(6 + 2\delta))$, $\sup_{n \in \mathbf{N}} |\mu^n f|/n^\alpha$ *is a weak* $(1,1)$ *operator.*

It is worth noticing that, when $\delta = 0$, α can be made smaller than $1/2$, which is the factor that, after Theorem 1.1.6 and Lemma 1.1.7, one would have expected to be the right normalization.

2. Convergence for non-centered measures

There are examples that show that when the measure μ does not have expected value zero, the whole sequence $\mu^n f(x)$ fails to converge even though the support of the measure generates $\mathbf{Z}$. The simplest example is $\mu = \frac{1}{2}(\delta_0 + \delta_1)$ (see [5]). However, even in such cases, there are convergence a.e. results along subsequences of powers, $\mu^{n_k} f(x)$. Bellow-Jones-Rosenblatt [2], [5] studied results of this type on L^p spaces. In [2], the authors show that if μ has finite support and is strictly aperiodic then $\lim_{n \to \infty} \mu^{2^{3^n}} f(x)$ exist almost everywhere for any $f \in L^p(X)$, $1 < p \leq \infty$. In this section, the analogous result for $p = 1$ is proven as well as an explicit criterion for lacunary sequences $\{n_k\}$ under which there is convergence a.e.

Let μ be a probability measure on the integers. Denote by ϕ_n a discrete version of the Gaussian distribution centered at 0 and normalized with the variance of μ, that is

$$\phi_n(k) = \begin{cases} \dfrac{1}{\sigma\sqrt{2\pi n}} \exp\left(-\dfrac{k^2}{2\sigma^2 n}\right) & \text{for } k \in [-n, n] \cap \mathbf{Z} \\ 0 & \text{otherwise} \end{cases}$$

where $\sigma = \sqrt{m_2(\mu)}$. Let $a = E(\mu)$.

PROPOSITION 2.1. *If μ is a strictly aperiodic probability measure on* $\mathbf{Z}$ *with* $m_{2+\delta}(\mu) < \infty$ *for some* $\delta > 0$ *and* $\{n_k\}_{k=0}^\infty$ *is a sequence in* $\mathbf{Z}^+$ *satisfying*

$$n_{k+1} > \gamma n_k \text{ for some } \gamma > 1, \tag{2.a}$$

then for any $f \in L^1(X)$, $\mu^{n_k} f(x)$ *converges a.e. if and only if* $\phi_{n_k} f(\tau^{a.n_k} x)$ *does.*

Proof. By calculations using Theorem 1.1.6 and the fact that $m_{2+\delta}(\mu) < \infty$, one can prove

$$\|\mu^n - \phi_n * \delta_{a.n}\|_1 \leq \frac{C}{n^\beta}$$

for some $\beta > 0$. Therefore, if a sequence satisfies (2.a), then

$$\sum_{k\geq 1} \|\mu^{n_k} - \phi_{n_k} * \delta_{a.n}\|_1 < \infty;$$

and consequently,

$$m\left\{x \in X / \sup_{k\in\mathbf{N}} \left|\mu^{n_k} f(x) - \phi_{n_k} f(\tau^{a.n_k} x)\right| > \lambda\right\} \leq C\frac{\|f\|_1}{\lambda}. \quad \square$$

In Bellow-Jones-Rosenblatt [5], the authors had already studied the behavior of $\phi_n f(x)$ along subsequences. Its convergence is related to the behavior of a maximal operator of block averages. We follow their notation. Let Ω be any subset of $\mathbf{Z} \times \mathbf{Z}^+$ and let

(2.b) $$\Omega_\alpha = \{(z, w) \in \mathbf{Z} \times \mathbf{Z}^+ / \text{ there exists } (s, n) \in \Omega \times \text{such that } |z - s| \leq \alpha(w - n)\}$$

be the cone of aperture α with vertices in Ω. Denote the cross section of Ω_α by

$$\Omega_\alpha(n) = \{z \in \mathbf{Z} / (z, n) \in \Omega_\alpha\}.$$

Consider now the maximal operator

$$M_\Omega f(x) = \sup_{(k,n)\in\Omega} \frac{1}{2n+1} \sum_{s=-n}^{n} f(\tau^{k+s}(s)) = \sup_{(k,n)\in\Omega} A_n f(\tau^k x)$$

There is a close relation between the weak type of the operator $M_\Omega f(x)$ and the growth of the cross sections $\Omega_\alpha(n)$.

THEOREM 2.2 (Bellow-Jones-Rosenblatt [5]). *The maximal operator $M_\Omega f$ is of weak type $(1, 1)$ if and only if there exists an $\alpha > 0$ for which the cross sections of Ω_α grow at most linearly; in other words, there is a positive constant $C < \infty$ such that $|\Omega_\alpha(n)| \leq Cn$ for all $n \geq 1$.*

Then, for the maximal operator with respect the Gaussian distributions, one has the following relation with the maximal operator on block averages.

THEOREM 2.3 (Bellow-Jones-Rosenblatt [5]). *The operator*

$$N_\Omega f(x) = \sup_{(k,n)\in\Omega} \phi_n f(\tau^k x)$$

is weak type $(1,1)$ *if and only if the operator* $M_{\overline{\Omega}}f$ *is of weak type* $(1,1)$, *where*

$$\overline{\Omega} = \left\{\left(k,\left[\sqrt{n}\,\right]\right):(k,n) \in \Omega\right\}.$$

From Proposition 2.1 and the above two theorems, it follows the following theorem.

THEOREM 2.4. *Let* μ *be a strictly aperiodic probability measure on* $\mathbf{Z}$ *with* $m_{2+\delta}(\mu) < \infty$ *for some* $\delta > 0$. *Let* τ *be an ergodic measure preserving transformation and* $\overline{\Omega} = \{(an_k, \sqrt{n_k})\}$, *where* $\{n_k\}$ *is an increasing sequence with* $n_k \geq \gamma n_{k-1}$ *for some* $\gamma > 1$. *If* $\overline{\Omega}_\alpha(n)$ *grows linearly for all* α, *then* $\lim_{k\to\infty} \mu^{n_k} f(x) = \int f\,dm$ *a.e. for all* $f \in L^1(X)$.

Proof. If the cross sections of $\overline{\Omega}$ grow linearly, then by Theorem 2.2, $M_{\overline{\Omega}}f$ is of weak type $(1,1)$. This implies, by Theorem 2.3, that $\sup_{(k,n)\in\Omega} \phi_n f(\tau^k x)$ is of weak type $(1,1)$. And finally, by Proposition 2.1, one obtains that $\sup_{k\in\mathbf{N}} \mu^{n_k} f$ is of weak type $(1,1)$. □

The next lemma characterizes the sequences with (2.a) for which the cross sections of Ω_α grow linearly.

LEMMA 2.5. *If* $\{n_k\}_{k\geq 1}$ *is a sequence with the growth condition* (2.a), *then the cross sections of* $\Omega = \{(an_k, \sqrt{n_k})/k \geq 1\}$ *grow linearly if and only if the function*

$$\Psi(\lambda) = \#\{n_k : \lambda < n_k \leq \lambda^2\}$$

is bounded.

Proof. Let $(n_k)_{k\geq 1}$ be a sequence with (2.a). Without loss of generality, one can assume $a = 1$.

First, assume that the cross sections of Ω grow linearly. It will then follow that Ψ is a bounded function. It suffices to consider cones with aperture $\alpha \leq \gamma - 1$. Then, for any $\lambda \in \mathbf{N}$, the cones with aperture α and vertices in $\Omega = \{(n_k, \sqrt{n_k})/k \geq 1\}$ at points with $n_k \geq \lambda$, have disjoint cross sections at level λ. Indeed, consider $\lambda < n_{k-1} < n_k \leq \lambda^2$. By condition (2.a), $n_k - n_{k-1} \geq (\gamma - 1)n_{k-1}$. Two consecutive cones are disjoint at level λ if

$$\begin{aligned} & n_{k-1} + \frac{\alpha}{2}\left(\lambda - \sqrt{n_{k-1}}\right) < n_k - \frac{\alpha}{2}\left(\lambda - \sqrt{n_k}\right) \\ \Leftrightarrow\ & \alpha\lambda - \frac{\alpha}{2}\left(\sqrt{n_k} + \sqrt{n_{k-1}}\right) < n_k - n_{k-1} \\ \Leftrightarrow\ & \alpha_\lambda < (\gamma - 1)n_{k-1} + \frac{\alpha}{2}\left(\sqrt{n_k} + \sqrt{n_{k-1}}\right). \end{aligned}$$

But $n_{k-1} > \lambda$ and $\alpha \le \gamma - 1$. Thus, the cross sections in consideration are disjoint. Now one can estimate the size of the whole cross section of $\Omega_\alpha(\lambda)$. The contribution of the cones with vertices corresponding to n_k's, $n_{k-1} \le \lambda$, does not exceed 3λ. And, by the above discussion, the remaining part is

$$2 \sum_{\lambda < n_k \le \lambda^2} \left(\lambda - \sqrt{n_k}\right).$$

Hence, $|\Omega_\alpha(\lambda)| \le C\lambda$ for some constant C, which gives

$$\sum_{\lambda < n_k \le \lambda^2} \left(\lambda - \sqrt{n_k}\right) \le C\lambda;$$

or equivalently,

$$\lambda \Psi(\lambda) - \sum_{\lambda < n_k \le \lambda^2} \sqrt{n_k} \le C\lambda,$$

where $\Psi(\lambda) = \#\{k : \lambda < n_k \le \lambda^2\}$. And then,

$$(*) \qquad \Psi(\lambda) \le C + \frac{1}{\lambda} \sum_{\lambda < n_k \le \lambda^2} \sqrt{n_k}$$

Let n_{j_0} be the first element of the sequence $\{n_k\}$ with $n_k > \lambda$. Then

$$\left\{n_k : \lambda < n_k \le \lambda^2\right\} = \left\{n_{j_0}, n_{j_0+1}, \cdots, n_{j_0+r}\right\},$$

where $r = \Psi(\lambda) - 1$. By the hypothesis on the sequence $\{n_k\}$,

$$n_{j0} < \left(\frac{1}{\gamma}\right)^r n_{j_0+r}, n_{j_0+1} < \left(\frac{1}{\gamma}\right)^{r-1} n_{j_0+r}, \cdots, n_{j_0+r-1} < \left(\frac{1}{\gamma}\right) n_{j_0+r}.$$

So the left hand side of $(*)$ is smaller than

$$C + \frac{1}{\lambda} \sum_{k=0}^{r} \left(\frac{1}{\sqrt{\gamma}}\right)^k \sqrt{n_{j_0+r}} \le C + \sum_{k=0}^{\infty} \left(\frac{1}{\sqrt{\gamma}}\right)^k = C' < \infty$$

because $\gamma > 1$ and $n_{j_0+r} < \lambda^2$.

The other direction is immediate. Indeed, if $\Psi(\lambda) \le M$ for all λ, then, for any $\alpha > 0$ and $\lambda > 0$,

$$\begin{aligned} |\Omega_\alpha(\lambda)| &\le (\text{contribution of cones with vertices} \le \lambda) + 2M\lambda \\ &\le (2 + 2M)\lambda. \quad \square \end{aligned}$$

COROLLARY 2.6. *If μ is a strictly aperiodic probability measure on* $\mathbf{Z}$, *then*

$$\lim_{n\to\infty} \mu^{2^{2^n}} f(x) \text{ and } \lim_{n\to\infty} \mu^{2^{2^n/n^p}} f(x)$$

exist a.e. for all $f \in L^1(X)$ and all $p \geq 0$; but

$$\lim_{n\to\infty} \mu^{2^{n^2}} f(x) \quad \text{and} \quad \lim_{n\to\infty} \mu^{2^{2\sqrt{n}}} f(x)$$

fail to exist on a set of positive measure, for some $f \in L^1(X)$

REFERENCES

1. L.E. BAUM and M. KATZ, *Convergence rates in the law of large numbers*, Trans. Amer. Math. Soc. **120** (1965), 108–123.
2. A. BELLOW, R. JONES and J. ROSENBLATT, . *Almost everywhere convergence of weighted averages*, preprint.
3. ______, *Almost everywhere convergence of convolution powers*, preprint.
4. ______, *Almost everywhere convergence of powers*, Academic Press, New York, 1989, pp. 99–120.
5. ______, *Convergence of moving averages*, Ergodic Theory Dynamical Systems **10** (1990), 43–62.
6. Y.S. CHOW and H. TEICHER, *Probability theory. Independence, interchangeability, martingales*, Springer–Verlag, New York, 1978.
7. K.L. CHUNG and P. ERDÖS, *Probability limit theorems assuming only the first moment I*, Mem. Amer. Math. Soc., no. 6, pp. 13–19.
8. S. FOGUEL, *On iterates of convolutions*, Proc. Amer. Math. Soc. **47** (1975), 368–370.
9. A. GARSIA, *Topics in almost everywhere convergence*, Lecture Notes in Advanced Math., vol. 4, Markham, Chicago, Illinois, 1970.
10. HUA LOO KENG, *Introduction to number theory*, Springer–Verlag, New York, 1982.
11. I.A. IBRAGIMOV and YU. V. LINNIK, *Independent and stationary sequences of random variables*, Wolters-Noordhoff (J.F.C. Kingman), Grontingen, 1971.
12. U. KRENGEL, *Ergodic theorems*, de Gruyter Studies in Mathematics, vol. 6 (Walter de Gruyter), New York, 1985.
13. D. ORNSTEIN, *A remark on the Birkhoff ergodic theorem*, Illinois J. Math. **15** (1971), 77–79.
14. V.I. OSELEDEC, *Markov chains, skew products and ergodic theorems for general dynamic systems*, Teor. Verojanost i Primenen **10** (1965), 551–557 (Russian); English transl. Annals of Prob. and Appl. **10** (1965), 499–504.
15. S. SAWYER, *Maximal inequalities of weak type*, Ann. of Math. **84** (1966), 157–174.

THE OHIO STATE UNIVERSITY
COLUMBUS, OHIO

STATE UNIVERSITY OF NEW YORK AT ALBANY
ALBANY, NEW YORK