Moment convergence and the law of iterated logarithm for additive functions

by

István Berkes (Graz) and Michel Weber (Strasbourg)

1. Introduction. Let $f$ be a strongly additive arithmetic function and set

\begin{equation}
A_n = \sum_{p \leq n} \frac{f(p)}{p}, \quad B_n = \left( \sum_{p \leq n} \frac{f^2(p)}{p} \right)^{1/2}.
\end{equation}

By a classical result of Erdős and Kac [7], if $|f(p)| = O(1)$ and $B_n \to \infty$, then we have

\begin{equation}
\lim_{N \to \infty} \frac{1}{N} \# \{ n \leq N : f(n) \leq A_N + xB_N \} = \left( \frac{2\pi}{x} \right)^{1/2} \int_{-\infty}^{x} e^{-u^2/2} \, du
\end{equation}

for all $x \in \mathbb{R}$. The same conclusion holds for unbounded $f(p)$, provided $f$ satisfies

\begin{equation}
\lim_{n \to \infty} \frac{1}{B_n^2} \sum_{p \leq n, |f(p)| \geq \varepsilon B_n} \frac{f^2(p)}{p} = 0 \quad \text{for any } \varepsilon > 0.
\end{equation}

(See Kubilius [13], Shapiro [17].) Condition (3) is the analogue of the Lindeberg condition for the central limit theorem in probability theory and the previous results show that the distributional behavior of additive functions is similar to that of sums of independent random variables. For extensions and further related results on the distribution of arithmetic functions see e.g. Kubilius [13], Elliott [4] and the references therein. Note that under mild technical conditions on $B_n$ the converse implication (2) $\Rightarrow$ (3) also holds (see Kubilius [13, p. 58]), but in general this converse is false (see Timofeev [18]).

The standard proofs of the central limit theorem (2) (and in fact of most results on the distributional behavior of additive functions) depend on

\[ 2000 \text{ Mathematics Subject Classification: Primary 11N60, 11N37; Secondary 60F05.} \]

Research of the first-named author supported by Hungarian National Foundation for Scientific Research, grants T 043037 and T 037886.
asymptotic estimates for the cardinality of the set
\[ \{ m \leq N : \alpha_{p_i}(m) = \alpha_i, i = 1, \ldots, s \} \]
where
\[ m = \prod_p p^{\alpha_p(m)} \]
is the prime factorization of \( m \) and \( 2 = p_1 < \cdots < p_s \) are the primes not exceeding \( r \), where \( r = r(N) \) satisfies \( \log r / \log N \to 0 \). Such estimates can be deduced using sieve methods and they show that “not too many” of the arithmetic functions \( \alpha_p \) are almost statistically independent with respect to the normalized counting measure on \( \{1, \ldots, N\} \).

A more elementary (although rather technical) proof was given by Halberstam [9] and simplified substantially by Billingsley [2], using the method of moments. They proved that letting
\[ F_N(t) = \frac{1}{N} \# \{ n \leq N : f(n) < a_N + tB_N \} \]
we have
\[ \lim_{N \to \infty} \int_{-\infty}^{\infty} t^r dF_N(t) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} t^r e^{-t^2/2} \, dt \quad (r = 1, 2, \ldots). \]  
From (4), the central limit theorem (2) follows immediately. The purpose of this paper is to show (see Theorem 2 below) that the \( r \)th moment on the left hand side of (4) is asymptotically equal to the \( r \)th moment of the standard Gaussian distribution not only for fixed \( r \), but also if \( r = r(N) \) tends to infinity not faster than \( \log \log B_N \). Just as the validity of (4) for all fixed \( r \) implies the central limit theorem (2), this generalized moment behavior will lead, via a simple analysis, to a law of the iterated logarithm for \( f(n) \).

In view of (2), it is natural to expect that under conditions similar to (3) the set
\[ H_t = \{ n : |f(n) - A_n| \geq t(2B_n^2 \log \log B_n)^{1/2} \} \]
is “large” for \( t < 1 \) and “small” for \( t > 1 \). However, no such result seems to exist in the literature. The reason is that ordinary asymptotic density of sequences of integers, used in the central limit theorem (2), is too crude to measure the set \( H_t \): the asymptotic density of \( H_t \) equals 0 for any \( t > 0 \), regardless of whether \( t > 1 \) or \( t < 1 \). In this paper we will show, however, that using a finer measure on subsets of \( \mathbb{N} \), depending on the growth of the variance function \( B_n \), there is a sharp difference between the cases \( t > 1 \) and \( t < 1 \) in (5).

Let \( \mu \) denote the measure on subsets of \( \mathbb{N} \) defined by
\[ \mu(\{1, \ldots, N\}) = \log^* B_N, \quad N = 1, 2, \ldots, \]
where the * means that we interpolate \( \log B_N \) linearly between the points \( 2^k, k = 0, 1, \ldots \). That is, \( \log^* B_N \) is linear in the intervals \( 2^k \leq N \leq 2^{k+1} \) \((k = 1, 2, \ldots)\) and coincides with \( \log B_N \) at the points \( N = 2^k \). We will prove the following

**THEOREM 1.** Assume that \( B_n \to \infty \) and

\[
|f(p)| = O(B_p^{1-\delta}) \quad \text{for some } \delta > 0.
\]

Then \( \mu(H_t) < \infty \) for \( t > 1 \) and \( \mu(H_t) = \infty \) for \( t < 1 \).

To clarify the meaning of Theorem 1, and in particular, of the measure \( \mu \), let \( X_p, p = 2, 3, 5, \ldots \), be independent random variables, defined on some probability space, such that \( X_p \) takes the values \( f(p) \) and 0 with probabilities \( 1/p \) and \( 1 - 1/p \), respectively. Let \( S_n = \sum_{p \leq n} X_p \). By the classical arithmetic theory (see e.g. Kubilius [13]), the sequence \( \{S_n, n \leq N\} \) is an almost exact probabilistic replica of \( \{f(n), n \leq N\} \), where the latter sequence is meant with respect to the normalized counting measure on \( \{1, \ldots, N\} \). Since under (7) the sequence \( X_p \) trivially satisfies the central limit theorem

\[
(S_n - A_n)/B_n \overset{D}{\to} N(0,1),
\]

this fact implies (2) and leads to a whole class of further interesting distribution results for additive functions. In contrast to this nice behavior, the probabilistic properties of the infinite sequences

\[\{f(n), n \geq 1\}, \quad \{S_n, n \geq 1\}\]

are in general quite different. For example, the central limit theorem (2) implies that the asymptotic density of the set \( G = \{n : f(n) > A_n\} \) is 1/2; on the other hand, the sequence \( X_p \) satisfies the Lindeberg condition (3) and thus by a general version of the arc sine law (see e.g. Prokhorov [16]) we have

\[
\frac{1}{N} \sum_{k \leq N} I(S_k > A_k) \overset{D}{\rightarrow} H
\]

where \( H \) is a nondegenerate probability distribution on \( \mathbb{R} \). The last relation obviously implies that the set \( \{n : S_n > A_n\} \) has no asymptotic density; actually, its lower density is 0 and upper density is 1 with probability 1. To remedy this trouble, introduce the logarithmic density

\[
\mu^*(A) = \lim_{N \to \infty} \frac{1}{\log B_N} \sum_{k \leq N, k \in A} \log(B_k/B_{k-1}), \quad A \subset \mathbb{N},
\]

and note that by the so-called almost sure central limit theorem (for a suitable version see Atlagh [1] or Ibragimov and Lifshits [11]) we have

\[
\mu^*\{n : (S_n - A_n)/B_n \leq x\} = (2\pi)^{-1/2} \int_{-\infty}^{x} e^{-u^2/2} du \quad \text{a.s.}
\]
for any $x \in \mathbb{R}$. This means that under $\mu^*$ the process $\{(S_n - A_n)/B_n, \; n \geq 1\}$ is ergodic: the relative time it spends in any interval $I$ equals the asymptotic probability of $I$. In particular, $\mu^*(\{n : S_n > A_n\}) = 1/2$ with probability one, i.e. using the measure $\mu^*$ the previously observed difference between the distributions of $\{f(n), \; n \geq 1\}$ and $\{S_n, \; n \geq 1\}$ disappears. This suggests that logarithmic measure is the natural one in studying statements in probabilistic number theory of “almost sure” type and Theorem 1 shows that it works for the law of the iterated logarithm.

The first law of the iterated logarithm for additive arithmetic functions was formulated, without proof, by Erdős in [5, Theorem VI]; see Hall and Tenenbaum [10, Chapter 1] for a proof. Erdős’ theorem involves ordinary asymptotic density and is very close in spirit to the classical LIL of probability theory. On the other hand, it concerns truncated sums $\sum_{p|n, \; p \leq u} f(p)$ as $u \to \infty$, an object different from $f(n)$. Specialized to the case $f(p) = 1$, Erdős’ theorem states that for any $\varepsilon > 0$ the asymptotic density of integers $m$ which have at least one divisor $d$ with

$$\omega(d) > \log \log d + (1 - \varepsilon) \sqrt{2 \log \log d \log_4 d}$$

is 1 and for every $\varepsilon > 0$ the density of integers $m$ having at least one divisor $d > A$ with

$$\omega(d) > \log \log d + (1 + \varepsilon) \sqrt{2 \log \log d \log_4 d}$$

is tending to 0 if $A \to \infty$. Here $\omega(n)$ denotes the number of different prime divisors of $n$ and $\log_r$ denotes $r$ times iterated logarithm.

Erdős’ theorem was later extended by Kubilius (see [13, Theorem 7.2]) and in several papers by Manstavičius (see [14] and the references therein). Using an extension of the concept “a.s. convergence”, applicable in the context of a sequence of probability spaces, Manstavičius gave a profound study of the LIL properties of additive functions, including e.g. refined Strassen type functional versions of the standard LIL (see [15]). Similarly to Erdős’ result, his results use truncated additive functions and ordinary asymptotic density. On comparison, our results involve logarithmic density, but pertain directly to the set (5).

The connection of the arithmetic central limit theorem (2) with almost sure central limit theory reveals a paradoxical property of additive functions from the probabilistic point of view. By the almost sure central limit theorem quoted above, the sequence $X_p$ satisfies

$$\lim_{N \to \infty} \frac{1}{\log B_N} \sum_{k \leq N} \log(B_k/B_{k-1})I\left\{\frac{S_k - A_k}{B_k} \leq x\right\} = \Phi(x) \quad \text{a.s.,}$$

where $\Phi$ denotes the standard Gaussian distribution function. Moreover, this relation fails if we replace logarithmic averages by ordinary averages. In
contrast, for additive functions $f(n)$ we have, by (2),
\[
\lim_{N \to \infty} \frac{1}{N} \sum_{k \leq N} I \left\{ \frac{f(k) - A_k}{B_k} \leq x \right\} = \Phi(x),
\]
and thus in this case the a.s. central limit theorem holds with ordinary averages. This shows that while the probabilistic behavior of additive functions is well understood in the case of distributional properties like the central limit theorem, much remains to be done in the case of “almost sure” type limit theorems.

Condition (7) obviously implies the Lindeberg condition (3); in fact, it even implies
\[
\lim_{n \to \infty} \frac{1}{B_n^{2+\delta}} \sum_{p \leq n} \frac{|f(p)|^{2+\delta}}{p} = 0,
\]
which is the analogue of the Lyapunov condition for the CLT in probability theory. In analogy with Kolmogorov’s classical condition (see [12]) for the law of the iterated logarithm, it is natural to expect that the LIL of our paper remains valid under
\[
f(p) = o(B_p/(\log \log B_p)^{1/2}).
\]
However, the methods of our paper are not strong enough to decide the validity of this conjecture.

We finally note that using deeper tools from probabilistic number theory based on sieve methods, Theorem 1 can be sharpened in the same way as so-called upper-lower class tests in probability theory improve the law of the iterated logarithm. (See e.g. Feller [8].) However, as our main interest in the present paper is the elementary moment approach, we do not investigate such improvements of Theorem 1 here.

2. Proofs. The first step of the argument is a truncation of the function $f$. Clearly $f = \sum_p f(p) \delta_p$, where $\delta_p$ is the indicator function of the set $\{n \in \mathbb{N} : p | n\}$. Let the function $f_n$ be defined by
\[
(8) \quad f_n = \sum_{p \leq \alpha_n} f(p) \delta_p
\]
where
\[
\alpha_n = n^{1/(\log \log B_n)^2}.
\]
Set further
\[
(9) \quad a_n = \sum_{p \leq \alpha_n} \frac{f(p)}{p}, \quad b_n = \left( \sum_{p \leq \alpha_n} \frac{f^2(p)}{p} \right)^{1/2}.
\]
Lemma 1. Let $X_p$, $p = 2, 3, 5, \ldots$, be independent random variables, defined on some probability space, such that $X_p$ takes the values $f(p)$ and 0 with probabilities $1/p$ and $1 - 1/p$, respectively. Let $S_n = \sum_{p \leq \alpha_n} X_p$. Then

$$E \left\{ \left( \frac{S_n - a_n}{b_n} \right)^{2r} \right\} \sim \mu_{2r} \quad \text{as} \quad n \to \infty, \quad \text{uniformly for} \quad 1 \leq r \leq 4 \log \log b_n,$$

where $\mu_{2r} = 1 \cdot 3 \cdot \ldots \cdot (2r - 1)$ is the $2r$th moment of the standard normal law.

Proof. Let

$$X'_p = X_p - \frac{f(p)}{p}, \quad S'_n = \sum_{p \leq \alpha_n} X'_p = S_n - a_n,$$

$$s_n^2 = ES_n^2 = \sum_{p \leq \alpha_n} f^2(p) \left( \frac{1 - 1}{p} \right).$$

By a recent result of Cuny and Weber on the speed of convergence of moments in the central limit theorem (see [3, Theorem 1.3]) we have

$$\left| E \left( \frac{|S'_n|}{s_n} \right)^{2r} - \mu_{2r} \right| \leq \left( C_1 \frac{r}{\log r} \right)^{2r} \max_{h \in \{1, 1/(2r-2)\}} \left( \sum_{p \leq \alpha_n} E|X'_p|^{2r} \right)^{h}$$

where $C_1$ is an absolute constant. Here $E|X'_p|^{2r} = E|X_p - EX_p|^{2r} \leq 2^{2r} E|X_p|^{2r}$ by Minkowski’s inequality and thus we get, using $|f(p)| \leq CB_1^{1-\delta}$,

$$\sum_{p \leq \alpha_n} E|X'_p|^{2r} \leq 2^{2r} \sum_{p \leq \alpha_n} \frac{|f(p)|^{2r}}{p}$$

$$\leq (2C)^{2r-2} B_n^{(2r-2)(1-\delta)} 4 \sum_{p \leq \alpha_n} \frac{f^2(p)}{p}$$

$$\leq 4(2C)^{2r-2} B_n^{2r-\delta(2r-2)}.$$

On the other hand, the well known relation

$$\sum_{p \leq n} \frac{1}{p} = \log \log n + O(1)$$

implies

$$\sum_{\alpha_n < p \leq n} \frac{1}{p} = \log \log n - \log \log \alpha_n + O(1) \leq 3 \log \log \log B_n \quad (n \geq n_0),$$

whence

$$B_n^2 - b_n^2 = \sum_{\alpha_n < p \leq n} \frac{f^2(p)}{p} \ll B_n^{2(1-\delta)} \sum_{\alpha_n < p \leq n} \frac{1}{p}$$

$$\ll B_n^{2(1-\delta)} \log \log \log B_n.$$
and thus $s_n^2 \sim b_n^2 \sim B_n^2$. The statement of the lemma now follows from (10), (11) and the fact that

$$\left( \frac{r}{\log r} \right)^{2r} \leq (4 \log \log b_n)^{8 \log \log b_n} \leq \exp\{(\log \log b_n)^2\} \leq b_n^{\delta/4} \leq B_n^{\delta/4}$$

for $1 \leq r \leq 4 \log \log b_n$, $n \geq n_0$.

Throughout what follows, $P_n$ denotes the normalized counting measure on $\{1, \ldots, n\}$ and $E_n$ denotes the corresponding expectation.

**Lemma 2.** We have

$$E_n\left\{ \left( f_n - a_n \right)^r b_n \right\} - E_n\left\{ \left( S_n - a_n \right)^r b_n \right\} \rightarrow 0$$

uniformly for $1 \leq r \leq 8 \log \log b_n$.

**Proof.** We follow Billingsley [2]. Clearly

$$E(S_n^r) = \sum_{u=1}^{r} \sum' \frac{r!}{r_1! \cdots r_u!} \sum'' E(X_{p_1}^{r_1} \cdots X_{p_u}^{r_u})$$

and, by (8),

$$E_n(f_n^r) = \sum_{u=1}^{r} \sum' \frac{r!}{r_1! \cdots r_u!} \sum'' E_n(Y_{p_1}^{r_1} \cdots Y_{p_u}^{r_u})$$

where $Y_p = f(p)\delta_p$, $\sum'$ extends over those $u$-tuples $(r_1, \ldots, r_u)$ of positive integers satisfying $r_1 + \cdots + r_u = r$ and $\sum''$ extends over those $u$-tuples $(p_1, \ldots, p_u)$ of primes satisfying $p_1 < \cdots < p_u \leq \alpha_n$. Clearly,

$$E(X_{p_1}^{r_1} \cdots X_{p_u}^{r_u}) = \frac{f(p_1)^{r_1} \cdots f(p_u)^{r_u}}{p_1 \cdots p_u}$$

and

$$E_n(Y_{p_1}^{r_1} \cdots Y_{p_u}^{r_u}) = \frac{1}{n} \left[ n \frac{1}{p_1 \cdots p_u} \right] f(p_1)^{r_1} \cdots f(p_u)^{r_u}.$$
Now
\[ E((S_n - a_n)^r) = \sum_{k=0}^{r} \binom{r}{k} E(S_n^k)(-a_n)^{r-k} \]
and \( E_n(f_n - a_n)^r \) has an analogous expansion. Comparing the two expansions term by term and applying (18) we get
\[
|E(S_n - a_n)^r - E_n(f_n - a_n)^r| \leq \sum_{k=0}^{r} \binom{r}{k} \frac{\alpha_n^k b_n^k}{n} |a_n|^{r-k} = \frac{1}{n} (\alpha_n b_n + |a_n|)^r
\]
\[ \leq \frac{1}{n} (2\alpha_n b_n)^r, \]
where we used
\[
|a_n| \leq \sum_{p \leq \alpha_n} |f(p)| \leq \left( \sum_{p \leq \alpha_n} \frac{f^2(p)}{p^2} \right)^{1/2} \alpha_n^{1/2} \leq b_n \alpha_n.
\]

Now
\[
B_{2n}^2 - B_n^2 = \sum_{n < p \leq 2n} \frac{f^2(p)}{p} \ll B_{2n}^{2(1-\delta)} \sum_{n < p \leq 2n} \frac{1}{p} \ll B_{2n}^{2(1-\delta)},
\]
which shows that \( B_{2n}/B_n \to 1 \) and thus \( B_n \) is slowly varying in the Karamata sense, which implies \( B_n \ll n^\varepsilon \) for any \( \varepsilon > 0 \). Thus for \( 1 \leq r \leq 8 \log \log b_n \) we have
\[
(2\alpha_n)^r \leq 2^8 \log \log B_n n^{8/\log \log B_n} \ll (\log B_n)^8 n^{1/2} = o(n)
\]
and Lemma 2 is proved.

We can now easily get

**Theorem 2.** We have
\[
E_n\left\{ \left( \frac{f - A_n}{B_n} \right)^{2r} \right\} \sim \mu_{2r} \quad \text{as } n \to \infty,
\]
uniformly for \( 1 \leq r \leq 4 \log \log B_n \).

**Proof.** By (13) we have \( b_n^2/B_n^2 = 1 + O(B_n^{-\delta}) \) and thus
\[
b_n/B_n = (1 + O(B_n^{-\delta}))^{r/2} = 1 + o(1)
\]
uniformly for \( 1 \leq r \leq 8 \log \log B_n \). Thus from Lemmas 1 and 2 it follows that
\[
E_n\left\{ \left( \frac{f_n - a_n}{B_n} \right)^r \right\} \sim E_n\left\{ \left( \frac{f_n - a_n}{b_n} \right)^r \right\} \sim \mu_r \quad \text{as } n \to \infty,
\]
uniformly for all even \( r \) with \( 1 \leq r \leq 8 \log \log B_n \). Now
\[
|f(m) - f_n(m)| \leq \sum_{\alpha_n < p \leq n} |f(p)| \delta_p(m), \quad 1 \leq m \leq n,
\]
and thus, similarly to the proof of Lemma 2, we have
\[ E_n|f - f_n|^r \leq \sum_{u=1}^{r} \sum' \frac{r!}{r_1! \cdots r_u!} \sum'' \frac{1}{n} \left[ \frac{n}{p_1 \cdots p_u} \right] |f(p_1)|^{r_1} \cdots |f(p_u)|^{r_u}, \]
where \( \sum' \) extends over those \( u \)-tuples \((r_1, \ldots, r_u)\) of positive integers satisfying \( r_1 + \cdots + r_u = r \) and \( \sum'' \) extends over those \( u \)-tuples \((p_1, \ldots, p_u)\) of primes satisfying \( \alpha_n < p_1 < \cdots < p_u \leq n \). Thus using (7) and (12) we get,
for \( n \geq n_0 \),
\[ E_n|f - f_n|^r \leq C^r B_n^{r(1-\delta)} \sum_{u=1}^{r} \sum' \frac{r!}{r_1! \cdots r_u!} \sum'' \frac{1}{p_1 \cdots p_u} \]
\[ = C^r B_n^{r(1-\delta)} \left( \sum_{\alpha_n < p \leq n} \frac{1}{p} \right)^r \]
\[ \leq (3C)^r B_n^{r(1-\delta)} (\log \log B_n)^r \leq (3C)^r B_n^{r(1-\delta/2)}, \]
where \( C \) is the constant implied by the \( O \) in (7). Thus letting \( \|g\|_{r,n} = E_n(|g|^r)^{1/r} \) for any arithmetic function \( g \), by Minkowski’s inequality we get
\[ E_n |f - a_n|/B_n \|_{r,n} = \| (f - a_n)/B_n \|_{r,n} \leq 3CB_n^{-\delta/2}. \]
Further by (7) and (12) we have
\[ |A_n - a_n| \leq CB_n^{1-\delta} \sum_{\alpha_n < p \leq n} \frac{1}{p} \leq CB_n^{1-\delta/2} \quad \text{for } n \geq n_0 \]
and thus replacing \( f - a_n \) by \( f - A_n \) in the first term on the left hand side of (20) results in a change \( \leq CB_n^{-\delta/2} \) of the norm.

Let now \( \varepsilon > 0 \). Relation (19) shows that for even \( r \) and \( n \geq n_0(\varepsilon) \) the second term on the left hand side of (20) lies in the interval \([((1 - \varepsilon)\mu_r)^{1/r}, ((1 + \varepsilon)\mu_r)^{1/r}]\) and thus
\[ \| (f - A_n)/B_n \|_{r,n} \leq ((1 + \varepsilon)\mu_r)^{1/r} + 4CB_n^{-\delta/2} \leq ((1 + 2\varepsilon)\mu_r)^{1/r}, \]
observing that \( \mu_r \geq 1 \) and
\[ (1 + 2\varepsilon)^{1/r} - (1 + \varepsilon)^{1/r} \geq 4CB_n^{-\delta/2} \]
for \( 1 \leq r \leq 8 \log \log B_n \) by the mean value theorem. A similar argument yields
\[ \| (f - A_n)/B_n \|_{r,n} \geq ((1 - 2\varepsilon)\mu_r)^{1/r} \]
and Theorem 2 is proved.

Using Theorem 2 we can now get upper and lower tail estimates for \( |f - A_n| \) using a method going back to Kolmogorov [12] in the context of the moment generating functions and to Erdős and Gál [6] in the case of moment convergence.
Lemma 3. We have
\[ P_n \{ |f - A_n| \geq (2(1 + \varepsilon)B_n^2 \log \log B_n)^{1/2} \} \ll \exp(-(1 + \varepsilon) \log \log B_n). \]

Proof. Let
\[ G(t) = P_n \{ |f - A_n| \geq (2tB_n^2 \log \log B_n)^{1/2} \}, \quad t > 0, \]
and
\[ Z_n = (f - A_n)^2/(2B_n^2 \log \log B_n). \]
Since
\[ \mu_{2r} = \frac{(2r)!}{2^r r!} \sim \sqrt{2} \left( \frac{2r}{e} \right)^r \quad \text{as } r \to \infty, \]
by Lemmas 1 and 2 we get, for \( 1 \leq r \leq 4 \log \log B_n, \ n \geq n_0, \)
\[ (21) \quad (r/e)^r (\log \log B_n)^{-r} \ll E_n Z_n^r \ll (r/e)^r (\log \log B_n)^{-r} \]
where the constants implied by \( \ll \) are absolute. By (21) and the Markov inequality
\[ (22) \quad G(t) = P_n(Z_n \geq t) \leq t^{-r} E_n Z_n^r \ll t^{-r} (r/e)^r (\log \log B_n)^{-r}. \]
If \( t \geq 3, \) we choose \( r = \lceil e \log \log B_n \rceil \) to get
\[ (23) \quad G(t) \ll t^{-2 \log \log B_n}, \quad t \geq 3. \]
For \( 0 < t < 3 \) we choose \( r = \lceil t \log \log B_n \rceil \) to get
\[ (24) \quad G(t) \ll \exp(-t \log \log B_n), \quad 0 < t < 3, \]
and Lemma 3 is proved.

Lemma 4. We have
\[ P_n \{ |f - A_n| \geq (2(1 - \varepsilon)B_n^2 \log \log B_n)^{1/2} \} \gg \exp(-(1 - \varepsilon^2/16) \log \log B_n). \]

Proof. Let
\[ D_1 = \{ 1 - \varepsilon \leq Z_n \leq 1 \}, \quad D_2 = \{ 0 \leq Z_n < 1 - \varepsilon \}, \]
\[ D_3 = \{ 1 < Z_n \leq 3 \}, \quad D_4 = \{ Z_n > 3 \}. \]
Then by (21) we have, for \( 1 \leq r \leq 4 \log \log B_n, \ n \geq n_0, \)
\[ (25) \quad G(1 - \varepsilon) = P_n(Z_n \geq 1 - \varepsilon) \geq P_n(D_1) \geq \int_{D_1} Z_n^r dP_n \]
\[ \geq A(r/e)^r (\log \log B_n)^{-r} - (I_2 + I_3 + I_4) \]
where \( A \) is an absolute constant and
\[ I_k = \int_{D_k} Z_n^r dP_n, \quad k = 2, 3, 4. \]
We choose \( r = \lceil (1 - \varepsilon/2) \log \log B_n \rceil \) and estimate \( I_2, I_3 \) and \( I_4 \) from above.
First we get, using (24) and \( G(t) = P_n(Z_n \geq t) \),
\[
I_2 = - \int_0^{1-\varepsilon} t^r dG(t) \leq 2r \int_0^{1-\varepsilon} t^{r-1} G(t) \, dt
\]
\[
\ll 2r \int_0^{1-\varepsilon} t^{r-1} \exp(-t \log \log B_n) \, dt
\]
\[
= 2r (\log \log B_n)^{-r} \int_0^{(1-\varepsilon) \log \log B_n} u^{r-1} e^{-u} \, du.
\]
Since \( u^{r-1} e^{-u} \) reaches its maximum at \( u = r - 1 \), which exceeds the upper limit of the last integral by the choice of \( r \), we get
\[
I_2 \leq 2r (1-\varepsilon)^r e^{-(1-\varepsilon) \log \log B_n}
\]
\[
\leq 4 \log \log B_n \cdot (1-\varepsilon)^{(1-\varepsilon)/2} \log \log B_n (\log B_n)^{-(1-\varepsilon)}
\]
\[
= 4 (\log \log B_n) (\log B_n)^{-\gamma}
\]
where
\[
\gamma = 1 - \varepsilon - (1 - \varepsilon/2) \log(1 - \varepsilon).
\]
Similarly,
\[
I_3 \leq 2r (\log \log B_n)^{-r} \int_{\log \log B_n}^{3 \log \log B_n} u^{r-1} e^{-u} \, du.
\]
Now the maximum of the integrand is reached at a point which is smaller than the lower limit of the integral and we get
\[
I_3 \leq 4 (\log \log B_n) (\log B_n)^{-1}.
\]
Finally, to estimate \( I_4 \) we proceed as with \( I_2 \), but instead of (24) we use (23) to get
\[
I_4 \leq 2r \int_3^{\infty} t^{r-1} G(t) \, dt \ll 2r \int_3^{\infty} t^{r-1} t^{-2 \log \log B_n} \, dt
\]
\[
\ll e^{-\log \log B_n} = (\log B_n)^{-1}.
\]
Now using \( r = [(1 - \varepsilon/2) \log \log B_n] \) we see that the first term on the right hand side of (25) is
\[
A(r/e)^r (\log \log B_n)^{-r} \gg (r/e)^r \left( \frac{r}{1 - \varepsilon/2} \right)^{-r} \gg (\log B_n)^{-\gamma'}
\]
where
\[
\gamma' = (1 - \varepsilon/2) - (1 - \varepsilon/2) \log(1 - \varepsilon/2)
\]
and the constants implied by \( \gg \) are absolute. Simple calculations show that for sufficiently small \( \varepsilon \) we have \( \gamma' < \gamma \) and \( \gamma' < 1 - \varepsilon^2/16 \), which implies that
all of $I_2$, $I_3$ and $I_4$ are of smaller order of magnitude than the expression in (28). Thus we get

$$G(1 - \varepsilon) \gg (\log B_n)^{-\gamma'} \gg (\log B_n)^{-(1 - \varepsilon^2/16)}$$

and Lemma 4 is proved.

We can now easily prove Theorem 1. Let $0 < \varepsilon < 1$. By Lemma 3 we have

$$P_{2k}\{|f - A_{2k}| \geq (2(1 + \varepsilon)B_{2k}^2 \log \log B_{2k})^{1/2}\} \ll \exp(-(1 + \varepsilon)\log \log B_{2k}).$$

As we have seen in the proof of Lemma 2, we have $B_{2k}/B_{2k-1} \to 1$. Also, the fluctuation of $A_n$ in the interval $[2^{k-1}, 2^k]$ is at most

$$\sum_{2^{k-1} < p \leq 2^k} \frac{|f(p)|}{p} \ll B_{2k}^{1-\delta} \sum_{2^{k-1} < p \leq 2^k} \frac{1}{p} \ll B_{2k}^{1-\delta}.$$

Thus the number of $j \in (2^{k-1}, 2^k]$ belonging to $H_{(1+2\varepsilon)^{1/2}}$ is

$$\ll 2^k \exp(-(1 + \varepsilon)\log \log B_{2k}) = \frac{2^k}{(\log B_{2k})^{1+\varepsilon}}.$$

By the definition of $\log^* B_N$, the $\mu$-measure of any point $j$ with $2^{k-1} < j \leq 2^k$ is

$$2^{-(k-1)} \log(B_{2k}/B_{2k-1}) \sim 2^{-(k-1)}(B_{2k}/B_{2k-1} - 1).$$

Thus

$$\mu(H_{(1+2\varepsilon)^{1/2}} \cap (2^{k-1}, 2^k]) \ll 2^{-k} \frac{B_{2k} - B_{2k-1}}{B_{2k-1}} \frac{2^k}{(\log B_{2k})^{1+\varepsilon}} \ll \int_{B_{2k-1}}^B \frac{1}{x(\log x)^{1+\varepsilon}} dx.$$

Summing over $k$ we get the first part of the theorem.

The proof of the second part is similar, but instead of (29) we use

$$P_{2k}^*\{|f - A_{2k}| \geq (2(1 - \varepsilon)B_{2k}^2 \log \log B_{2k})^{1/2}\} \gg \exp(-(1 - \varepsilon^2/16)\log \log B_{2k})$$

where $P_{2k}^*$ denotes uniform probability on the set $\{2^{k-1}+1, \ldots, 2^k\}$. Relation (30) is similar to our lower tail estimate

$$P_{2k}\{|f - A_{2k}| \geq (2(1 - \varepsilon)B_{2k}^2 \log \log B_{2k})^{1/2}\} \gg \exp(-(1 - \varepsilon^2/16)\log \log B_{2k})$$

in Lemma 4 and can be proved in the same way.

**Acknowledgments.** We thank the referee for valuable comments.
References


Received on 30.8.2005 and in revised form on 13.1.2006 (5060)