TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 269, Number 2, February 1982

PROBABILISTIC AND DETERMINISTIC AVERAGING

BY

N. H. BINGHAM AND CHARLES M. GOLDIE

ABSTRACT. Let $\{S_n\}$ be a random walk whose step distribution has positive mean μ and an absolutely continuous component. For any bounded measurable function f, a Marcinkiewicz-Zygmund strong law in an r-quick version (a 'Lai strong law') is proved for $f(S_n)$, assuming existence of a suitable higher moment of the step distribution. This is extended to show $n^{-\alpha}\{\Sigma_1^n f(S_k) - \int_0^n f(\mu t)\,dt\} \to 0$ (r-quickly). These results remain true when the step distribution is lattice, provided f is constant between lattice points. Certain intermediate results on renewal theory, mixing, local limit theory, ladder height, and a strong law of Lai for mixing random variables are of independent interest.

1. Introduction. Let $X_1, X_2, \ldots$ be independent identically distributed random variables on some probability space $(\Omega, \mathcal{F}, P)$, and write F, μ, σ^2 for the distribution function, mean and variance of X_1. Setting $S_n = X_1 + \cdots + X_n$, $S_0 = 0$, we write $\mathcal{P}$, $\mathcal{P}_n$ for the laws of X_1, S_n. One of the following assumptions, familiar in the context of local limit theorems and decompositions of renewal measures, will always be in force:

(LLT-I) Some $\mathcal{P}_k$ has an absolutely continuous component.

(LLT-I′) $\mathcal{P}$ has an absolutely continuous component.

(LLT-II) $\mathcal{P}$ is supported by the integers $\mathbf{Z}$, and is not supported by any ideal $k\mathbf{Z}$, $k > 1$.

Our results in the LLT-II case may be re-expressed for any lattice law $\mathcal{P}$, by suitable change of location and scale.

Throughout, let f be a real-valued bounded measurable function: $\sup|f(\cdot)| < M$. Our concern is to link the limiting behaviour of f and $f(S_n)$. When LLT-II is in force (S_n supported by the integers), clearly $f(S_n)$ can be compared only with values of f at integer points, and so we assume for that case without further comment that f is constant on each interval $[k, k+1)$, $k \in \mathbf{Z}$. Using that convention, the strong law for $f(S_n)$ given by Meilijson [**30**] may be stated as follows.

THEOREM A. *Assume* $0 < \mu < \infty$. *Assume* LLT-I *or* LLT-II. *Then*

$$n^{-1}\left\{\sum_{k=1}^{n} f(S_k) - \int_0^n f(\mu x)\,dx\right\} \to 0 \quad a.s. \tag{1.1}$$

Received by the editors November 25, 1980.

1980 *Mathematics Subject Classification*. Primary 60F15, 60F05, 40G05.

Key words and phrases. Backward recurrence time, Cesàro convergence, ladder height, laws of large numbers, local limit theorem, martingale, renewal theory, r-quick, strong mixing, variation norm.

© 1982 American Mathematical Society
0002-9947/82/0000-0766/$08.00

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

(In fact [**30**] assumes LLT-I′ rather than LLT-I, but the proof extends to LLT-I without trouble. For an alternative proof of the lattice case see Berbee [**3**, Proposition 6.2].)

In this paper we extend Theorem A under extra moment conditions, proving a Marcinkiewicz-Zygmund law by replacing n^{-1} by $n^{-\alpha}$, for suitable $\alpha > \frac{1}{2}$, in (1.1), and strengthening the mode of convergence to the 'r-quick' convergence of Strassen [**44**]. Our results are thus 'r-quick versions of Marcinkiewicz-Zygmund strong laws', in the terminology of Lai's important paper [**27**], one of whose results we extend. In honour of [**27**] we put forward 'Lai strong laws' as an appropriate and more concise label.

Recall ([**44**], [**26**]) that a sequence of random variables U_n is said to tend to zero *r-quickly* ($r > 0$) if for each $\varepsilon > 0$ the random variable $\sup\{n \geqslant 1: |U_n| \geqslant \varepsilon\}$ has finite rth moment (by convention, $\sup \varnothing = 0$). Such convergence implies a.s. convergence. Define $x^+ := \max(x, 0)$, $x^- := -\min(x, 0)$.

THEOREM 1 (LAI STRONG LAW). *Assume*

(i) LLT-I′ *or* LLT-II *holds*;

(ii) $\mu > 0$, $E\{(X_1^+)^l\} < \infty$ *for some* $l > 2$, $E\{(X_1^-)^\lambda\} < \infty$ *for some* $\lambda \geqslant 2$.

Choose $\alpha > \max(\frac{1}{2}, (l-1)^{-1})$, *and then* r *such that*

$$0 < r < \min\left\{\left(\alpha - \tfrac{1}{2}\right)\lambda, \alpha(l-1) - 1\right\}.$$

Then

$$n^{-\alpha} \sum_{k=1}^{n} \{f(S_k) - Ef(S_k)\} \to 0, \quad r\text{-quickly}. \tag{1.2}$$

Note the trade-off between the values of α and r; the smaller α, the stronger the statement that $n^{-\alpha}\sum_1^n \{f(S_k) - Ef(S_k)\} \to 0$, but then the smaller r, and the weaker the mode of convergence. The value $\alpha = 1$ is always available. In the nonlattice case the slight extra restriction of LLT-I′ is needed for technical reasons (see §2.7). When $F(0+) = 0$, that is, $X_1 > 0$ a.s., condition LLT-I suffices.

THEOREM 2 (CONVERGENCE OF FIRST MOMENTS). *Assume*

(i) LLT-I *or* LLT-II;

(ii) $\mu > 0$, $E\{|X_1|^{2+\delta}\} < \infty$ *for some* δ *satisfying* $0 < \delta \leqslant 1$.

Then for each $\varepsilon > 0$,

$$n^{-1} \sum_{k=1}^{n} Ef(S_k) - n^{-1} \int_0^n f(\mu x)\, dx = o(n^{-\delta/2+\varepsilon}), \qquad n \to \infty. \tag{1.3}$$

Combining Theorems 1 and 2 we have

THEOREM 2′ (r-QUICK EQUICONVERGENCE). *Assume*

(i) LLT-I′ *or* LLT-II;

(ii) $\mu > 0$, *and* $E\{(X_1^+)^l\} < \infty$, $E\{(X_1^-)^\lambda\} < \infty$ *for some* $l > 2$, $\lambda > 2$.

Choose α *such that* $\alpha > \max(\frac{1}{2}, 2 - \frac{1}{2}l, 2 - \frac{1}{2}\lambda)$ *and then* r *such that* $0 < r < \min\{(\alpha - \frac{1}{2})\lambda, \alpha(l-1) - 1\}$. *Then*

$$n^{-\alpha}\left\{\sum_{k=1}^{n} f(S_k) - \int_0^n f(\mu x)\, dx\right\} \to 0, \quad r\text{-quickly}. \tag{1.4}$$

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

In the important special case when $E(|X_1|^3) < \infty$ the parameter ranges in Theorems 1 and 2′ reduce to $\alpha > \frac{1}{2}$, $0 < r < 2\alpha - 1$.

We write

(I) $f(x) \to c$ (C_1) as $x \to \infty$

for the assertion that f converges to c under the Cesàro summability method of order 1. Probabilistic analogues of (I) are

(II) $n^{-1}\Sigma_1^n Ef(S_k) \to c$;

(III) $n^{-1}\Sigma_1^n f(S_k) \to c$ in probability;

(IV) $n^{-1}\Sigma_1^n f(S_k) \to c$ a.s.;

(V) $n^{-1}\Sigma_1^n f(S_k) \to c$, r-quickly.

As f is bounded, I is equivalent to

$$n^{-1}\int_0^n f(\mu x)\,dx = (n\mu)^{-1}\int_0^{n\mu} f(y)\,dy \to c \qquad (n \in \mathbf{Z}, \to \infty),$$

and hence (I) is equivalent to (IV) by Theorem A, if the conditions hold. Integrating (1.1) gives the equivalence of (I) and (II), and then (III) is equivalent to these assertions because (IV) implies (III) implies (II). We are led to the following corollary of Theorems A and 1.

COROLLARY 1.1. *Assume* (i) LLT-I′ *or* LLT-II; (ii) $0 < \mu < \infty$. *Then* (I), (II), (III) *and* (IV) *are equivalent. If, in addition,* $E(|X_1|^{2+\delta}) < \infty$ *for some* $\delta > 0$ *then, for every* r *satisfying* $0 < r < \min(1 + \frac{1}{2}\delta, \delta)$, (V) *is equivalent to* (I)–(IV).

The behaviour of $f(S_n)$ when $\mu = 0$ or when no assumption is made about the existence of μ can be quite different from the above. For the latter case see Meilijson [**30**, Result 2]. The $\mu = 0$ case is reviewed in §5.6 below. Pursuing the $\mu > 0$ case further, it is natural to ask whether versions of the CLT (central limit theorem) and LIL (law of the iterated logarithm) can be obtained. We leave these questions open in the general case. In the special case $\mathcal{P}$ exponential ($\{S_n\}$ Poisson) we give some results of this type below, but even here the results are less clearcut than those above.

The proof of Theorem 1, which is lengthy, follows in §2; this is divided into §2.1–§2.8 for convenience. The proofs of Theorems 2 and 2′ constitute §3. The CLT and LIL for the Poisson case are discussed in §4. Various complements are in §5.

The other principal results are

Theorem 3a, b (r-quick convergence in renewal theory),

Theorem 4 (strong mixing of backward recurrence times),

Theorem 5 (Lai strong law for uniformly bounded strong-mixing random variables),

Theorem 6 (absolute continuity of distribution of ladder height).

All these are in §2. In §3 the intermediate results are local limit theorems in variation-norm terms, which may also be of independent interest.

A word on format: whereas theorems have global (consecutive) numberings as above, lemmas, propositions, formulas etc. have local (decimal) numbering, by section or subsection.

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

2. Proof of Theorem 1.

2.1. *Outline and notation.* Because of the length of the proof, we pause to describe the method.

(a) In §2.8 we use ladder methods to reduce the general case to the 'renewal-theoretic' case where $F(0+) = 0$ ($X_n > 0$ a.s.). When $\mathcal{P}$ has an absolutely continuous component we need the ladder-height distributions to have the same property (§2.7). To preserve r-quickness, we use a result of Gut [17] on finiteness of moments of ladder-heights and epochs, and a Lai strong law for certain martingale difference sequences (§2.2).

(b) §§2.3–2.6 are aimed at proving the special case of Theorem 1 in which $F(0+) = 0$. Under this assumption, S_n increases and $N_t := \max\{n: S_n \leqslant t\}$ is a renewal process. Write $Y_k := \sum_{N_{k-1}<j\leqslant N_k} f(S_j)$. Then the results of §2.3 on r-quick convergence for renewal processes are used to show (1.1) equivalent to

$$n^{-\alpha}\sum_1^n (Y_k - EY_k) \to 0, \quad r\text{-quickly}.$$

Setting $\mathcal{F}_t = \sigma\{N_u: 0 \leqslant u \leqslant t\}$, the results of §2.2 show that this is equivalent to

$$n^{-\alpha}\sum_1^n \{E(Y_k|\mathcal{F}_{k-1}) - EY_k\} \to 0, \quad r\text{-quickly}.$$

Let $B_t := t - S(N_t)$ be the spent lifetime (backward recurrence time) at time $t \geqslant 0$. We show that the above can be rewritten as $n^{-\alpha}\sum_1^n \{g_k(B_{k-1}) - Eg_k(B_{k-1})\} \to 0$ r-quickly for certain uniformly bounded nonrandom functions g_k.

(c) In §2.4 we prove that $\{B_k\}$ is strong-mixing (in the sense of Rosenblatt [37]), whence so is $\{g_k(B_{k-1})\}$. In §2.5 we prove a Lai strong law for uniformly bounded strong-mixing (but not necessarily stationary) sequences of random variables. Combining these results, the proof when $F(0+) = 0$ is completed in §2.6.

(d) We note that $\{S_n\}$ itself cannot be strong-mixing, a counterexample being given in §5.2. It is this which necessitates the renewal-theoretic route outlined above, so as to be able to exploit the key Theorem 5 (below).

To avoid complicated suffices we shall, for a stochastic process Z, change at will between notations such as $Z(t)$, Z_t.

2.2. *Lemmas on r-quick convergence.*

LEMMA 2.2.1. *If $U_n \to 0$, r-quickly, and $U_n' \to 0$, r'-quickly, then $U_n + U_n' \to 0$, $(r \wedge r')$-quickly.*

We leave the proof as an exercise.

LEMMA 2.2.2. *For $r > 0$, $U_n \to 0$, r-quickly if and only if*

$$\int_0^\infty t^{r-1}P\Big(\sup_{k\geqslant t} |U_k| \geqslant \varepsilon\Big)\, dt < \infty \quad \textit{for all } \varepsilon > 0.$$

PROOF. See Lai [27, Lemma 4], Chow and Lai [9, p. 63].

LEMMA 2.2.3. *Assume the martingale difference sequence $\{\xi_n\}$ satisfies $E(|\xi_n|^\nu) \leqslant M$, $n = 1, 2, \ldots,$ where M and $\nu \geqslant 2$ are constants. Then for every α, r satisfying $\alpha > \frac{1}{2}$, $0 < r < (\alpha - \frac{1}{2})\nu$, $n^{-\alpha}\sum_1^n \xi_k \to 0$, r-quickly.*

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

PROOF. For some constant $K = K_\nu$,

$$E\left(\left|\sum_1^n \xi_k\right|^\nu\right) \leqslant Kn^{\nu/2}, \qquad n = 1, 2, \ldots,$$

by orthogonality for $\nu = 2$ and by [**43**, Theorem 3.7.8(i)] for $\nu > 2$. Then

$$P\left(\max_{k=1,\ldots,n}\left|\sum_1^k \xi_j\right| \geqslant \varepsilon n^\alpha\right) = P\left(\max_{k=1,\ldots,n}\left|\sum_1^k \xi_j\right|^\nu \geqslant (\varepsilon n^\alpha)^\nu\right)$$
$$\leqslant (\varepsilon n^\alpha)^{-\nu} E\left(\left|\sum_1^n \xi_j\right|^\nu\right)$$

by Doob's inequality [**15**, p. 314], since $|\sum_1^n \xi_j|^\nu$ is a nonnegative L_1 submartingale. Thus

$$\sum_{n=1}^\infty n^{r-1} P\left(\max_{k=1,\ldots,n}\left|\sum_1^k \xi_j\right| \geqslant \varepsilon n^\alpha\right) \leqslant \sum_1^\infty n^{r-1}(\varepsilon n^\alpha)^{-\nu} K n^{\nu/2} < \infty.$$

The result follows by [**27**, Lemma 4].

2.3. *Renewal theory*. Assume $F(0+) = 0$ throughout this subsection. Defining N_t as in §2.1, and extending the definition of r-quick convergence to a continuous time parameter in the obvious way, we consider the r-quick convergence to zero of $t^{-\alpha}N_t$, for $\alpha > 1$, or $t^{-\alpha}(N_t - t/\mu)$, for $\alpha \leqslant 1$. In the former case much more than r-quick convergence is true, without conditions:

THEOREM 3a. *For each $\alpha > 1$, $\varepsilon > 0$ the random variable*

$$T_{\alpha,\varepsilon} = \sup\{t > 0 \colon t^{-\alpha}N_t > \varepsilon\}$$

has tail behaviour

$$P(T_{\alpha,\varepsilon} \geqslant t) \leqslant ae^{-bt^\alpha}, \quad \text{all } t \geqslant 0, \tag{2.3.1}$$

for some positive constants $a(\alpha, \varepsilon)$, $b(\alpha, \varepsilon)$. Hence $t^{-\alpha}N_t \to 0$ r-quickly for every r.

PROOF. Let $[\cdot]$ denote integer part. On the event $\{T_{\alpha,\varepsilon} \geqslant t\}$ we have $N_u > \varepsilon u^\alpha$ for some $u \geqslant t$; hence $S([\varepsilon u^\alpha] + 1) \leqslant u$ and so $S_k \leqslant (k/\varepsilon)^{1/\alpha}$ where $k = [\varepsilon u^\alpha] + 1 > \varepsilon t^\alpha$. We truncate as follows. Define $m = \frac{1}{2}E(X_1 \wedge 1)$, $X_n' = 2m - (X_n \wedge 1)$, $S_n' = X_1' + \cdots + X_n'$. Then

$$2nm - S_n' = \sum_1^n \min(X_j, 1) \leqslant S_n, \qquad n = 1, 2, \ldots.$$

Hence, still on the event $\{T_{\alpha,\varepsilon} \geqslant t\}$, we have for some $k > \varepsilon t^\alpha$ that $S_k' \geqslant 2mk - (k/\varepsilon)^{1/\alpha}$, and the latter will be at least km provided $t > (\varepsilon m)^{-1/(\alpha-1)}$. So for all such t,

$$P(T_{\alpha,\varepsilon} \geqslant t) \leqslant P(S_k' \geqslant mk \text{ for some } k > \varepsilon t^\alpha) \leqslant \sum_{k > \varepsilon t^\alpha} P(S_k' \geqslant mk). \tag{2.3.2}$$

Because the X_k' are i.i.d. with $|X_k'| \leqslant 1$, $EX_k' = 0$, we may use Prohorov's exponential bound. Set $v = \operatorname{var} X_1'$. By [**43**, Theorem 5.2.2(ii)],

$$P(S_n' \geqslant mn) \leqslant \exp\left\{-\tfrac{1}{2}mn \operatorname{arcsinh}\left(\tfrac{1}{2}m/v\right)\right\} = e^{-Kn},$$

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

where $K > 0$ since $m > 0$. From (2.3.2),

$$P(T_{\alpha,\varepsilon} \geqslant t) \leqslant (1 - e^{-K})^{-1}\exp\{-K([\varepsilon t^{\alpha}] + 1)\}$$
$$\leqslant (1 - e^{-K})^{-1}\exp(-K\varepsilon t^{\alpha}),$$

as required. The constant a can be adjusted so that (2.3.1) applies also for $t \leqslant (\varepsilon m)^{-1/(\alpha-1)}$.

THEOREM 3b. *Let m, α, l be any constants satisfying $\frac{1}{2} < \alpha \leqslant 1$, $l\alpha > 1$, $m > 0$. Then*

$$t^{-\alpha}(N_t - t/m) \to 0, \quad (l\alpha - 1)\text{-quickly as } t \to \infty \tag{2.3.3}$$

if and only if

$$E(X_1^l) < \infty \quad \text{and} \quad EX_1 = m. \tag{2.3.4}$$

PROOF. By Lemma 2.2.2, (2.3.3) is equivalent to

$$\int_0^\infty u^{l\alpha-2}P\Big(\sup_{t \geqslant u} t^{-\alpha}|N_t - t/m| \geqslant \varepsilon\Big)\,du < \infty, \quad \text{all } \varepsilon > 0. \tag{2.3.5}$$

By Theorem 3 of Baum and Katz [1], (2.3.4) is equivalent to

$$\int_0^\infty u^{l\alpha-2}P\Big(\sup_{k \geqslant u} k^{-\alpha}|S_k - km| \geqslant \varepsilon\Big)\,du < \infty, \quad \text{all } \varepsilon > 0. \tag{2.3.6}$$

By suitable rescaling, we may assume without loss of generality that m in (2.3.5) and (2.3.6) is 1, lattice considerations being irrelevant here. To prove (2.3.5) and (2.3.6) equivalent, the same proof works in each direction. This is because $\{N_t\}$ and $\{S_n\}$ are essentially inverses of each other: $S_n = \sup\{t\colon N_t < n\}$, $N_t = \max\{n\colon S_n \leqslant t\}$, and there are only minor complications because of the step-function nature of the sample paths. We therefore prove only that (2.3.6) implies (2.3.5).

Fix $\varepsilon > 0$. Then

$$\begin{aligned}\int_0^\infty u^{l\alpha-2}P\Big(\sup_{t \geqslant u} t^{-\alpha}|N_t - t| \geqslant \varepsilon\Big)\,du \\ \leqslant \int_0^\infty u^{l\alpha-2}P\Big(\sup_{t \geqslant u} |t^{-1}N_t - 1| > \tfrac{1}{2}\Big)\,du \\ + \int_0^\infty u^{l\alpha-2}P\Big(\sup_{t \geqslant u} t^{-\alpha}|N_t - t| \geqslant \varepsilon,\ \sup_{t \geqslant u} |t^{-1}N_t - 1| \leqslant \tfrac{1}{2}\Big)\,du.\end{aligned} \tag{2.37}$$

If $t^{-1}N_t > 3/2$ then $S([3t/2] + 1) \leqslant t$ and so $S_k \leqslant 2k/3$ where $k = [3t/2] + 1 > 3u/2$. On the other hand if $t^{-1}N_t < \frac{1}{2}$ then $S_{[t/2]+1} > t$ and so $S_k > 2(k - 1)$ where $k = [\frac{1}{2}t] + 1 > \frac{1}{2}u$. In this case we further have $S_k > 4k/3$ so long as $u \geqslant 6$. So

$$\begin{aligned}\int_0^\infty u^{l\alpha-2}P\Big(\sup_{t \geqslant u} |t^{-1}N_t - 1| > \tfrac{1}{2}\Big)\,du \\ \leqslant \int_0^6 u^{l\alpha-2}\,du + \int_0^\infty u^{l\alpha-2}P\Big(\sup_{k > u/2} k^{-1}|S_k - k| \geqslant \tfrac{1}{3}\Big)\,du \\ \leqslant 6^{l\alpha-1}/(l\alpha - 1) + 2^{l\alpha-1}\int_0^\infty u^{l\alpha-2}P\Big(\sup_{k \geqslant u} k^{-1}|S_k - k| \geqslant \tfrac{1}{3}\Big)\,du \\ \leqslant 6^{l\alpha-1}/(l\alpha - 1) + 2^{l\alpha-1}\int_0^\infty u^{l\alpha-2}P\Big(\sup_{k \geqslant u} k^{-\alpha}|S_k - k| \geqslant \tfrac{1}{3}\Big)\,du < \infty\end{aligned}$$

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

by (2.3.6) (with $m = 1$). As for the other term on the right of (2.3.7), if $v^{-\alpha}|N_v - v| \geqslant \varepsilon$ for some $v \geqslant u$ and $|t^{-1}N_t - 1| \leqslant \frac{1}{2}$ for all $t \geqslant u$ then $|N_v - v| \geqslant \varepsilon(\frac{2}{3})^\alpha N_v^\alpha$ and $N_u \geqslant \frac{1}{2}u$. Setting $k = N_v$ we find $S_k \leqslant v < S_{k+1}$ so that one of the inequalities

$$|k - S_k| \geqslant \varepsilon\left(\tfrac{2}{3}\right)^\alpha k^\alpha, \qquad |k - S_{k+1}| \geqslant \varepsilon\left(\tfrac{2}{3}\right)^\alpha k^\alpha$$

must hold. Also $k \geqslant N_u \geqslant \frac{1}{2}u$. Thus

$$\begin{aligned}
&\int_0^\infty u^{l\alpha-2}P\left(\sup_{t\geqslant u} t^{-\alpha}|N_t - t| \geqslant \varepsilon, \sup_{t\geqslant u} |t^{-1}N_t - 1| \leqslant \tfrac{1}{2}\right) du \\
&\qquad \leqslant \int_0^\infty u^{l\alpha-2}P\left(\sup_{k\geqslant u/2} k^{-\alpha}|k - S_k| \geqslant \varepsilon\left(\tfrac{2}{3}\right)^\alpha\right) du \\
&\qquad\quad + \int_0^\infty u^{l\alpha-2}P\left(\sup_{k\geqslant u/2} k^{-\alpha}|k - S_{k+1}| \geqslant \varepsilon\left(\tfrac{2}{3}\right)^\alpha\right) du \\
&\qquad = 2^{l\alpha-1}\int_0^\infty u^{l\alpha-2}P\left(\sup_{k\geqslant u} k^{-1}|k - S_k| \geqslant \varepsilon\left(\tfrac{2}{3}\right)^\alpha\right) du \\
&\qquad\quad + 2^{l\alpha-1}\int_0^\infty u^{l\alpha-2}P\left(\sup_{k\geqslant u} k^{-\alpha}|k - S_{k+1}| \geqslant \varepsilon\left(\tfrac{2}{3}\right)^\alpha\right) du.
\end{aligned}$$

By (2.3.6) the first term on the right above is finite, and the second term also by trivial modifications. Thus everything in (2.3.7) is finite, which gives (2.3.5).

2.4. *Strong mixing of backward recurrence times.* Again $F(0+) = 0$ is assumed throughout this subsection. Analogously to the backward recurrence time $B_t := t - S(N_t)$ we write $F_t := S(N_t + 1) - t$ $(t \geqslant 0)$ for the forward recurrence time or remaining lifetime.

A sequence of random variables $\{U_n\}$ is called (Rosenblatt) *strong mixing* with *mixing coefficients* $\rho_n \downarrow 0$ if for $m, n = 0, 1, \ldots, C \in \sigma(U_0, \ldots, U_m)$, $D \in \sigma(U_{m+n}, U_{m+n+1}, \ldots)$,

$$|P(C \cap D) - P(C)P(D)| \leqslant \rho_n. \tag{2.4.1}$$

THEOREM 4. *Assume* LLT-I *or* LLT-II, *and suppose* $E(X_1^l) < \infty$ *for some* $l \geqslant 2$. *Then* $\{B_k\}$ *is strong mixing, with mixing coefficients* $\rho_n = O(n^{-(l-1)})$.

Our proof depends on the following lemma.

LEMMA 2.4.1. *Under the conditions of Theorem* 4, *there exists a constant* c *such that, for all* $m, n = 0, 1, \ldots,$ $D \in \sigma(B_{m+n}, B_{m+n+1}, \ldots)$, $x \in [0, \frac{1}{2}n]$ (*with* x *an integer in the lattice case*),

$$|P(D|F_m = x) - P(D)| \leqslant cn^{-(l-1)}. \tag{2.4.2}$$

PROOF. First assume LLT-I. Write $\mathfrak{B}$, $\mathfrak{B}^\infty$ for the Borel σ-fields in $\mathbf{R}$, $\mathbf{R}^\infty$. Then

$$D = \{\omega\colon (B_{m+n}(\omega), B_{m+n+1}(\omega), \ldots) \in D^0\} \tag{2.4.3}$$

for some $D^0 \in \mathfrak{B}^\infty$. The transition probability

$$Q(y, E) := P((B_{m+n+1}, B_{m+n+2}, \ldots) \in E|B_{m+n} = y), \qquad y \geqslant 0, E \in \mathfrak{B}^\infty, \tag{2.4.4}$$

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

is well defined. Write $U = \sum_0^\infty F^{(k)}$ for the renewal function, $\bar{F}(x) := 1 - F(x)$ for the tail of F. Given $F_m = x < n$, the conditional distribution of B_{m+n} is just the unconditional distribution of B_{n-x}; hence (see e.g. [**35**, p. 168])

$$(2.4.5)\quad P(B_{m+n} \in E | F_m = x) = \int_E \bar{F}(y) U(n - x - dy), \qquad E \in \mathfrak{B}, x < n.$$

The sequence $\{B_0, \ldots, B_m, F_m, B_{m+n}, B_{m+n+1}, \ldots\}$ is Markovian, whence

$$(2.4.6)\quad P(D | F_m = x) = \int\int_{\{(y, \mathbf{z}) \in D^0\}} Q(y, d\mathbf{z}) \bar{F}(y) U(n - x - dy), \qquad x < n.$$

We define

$$(2.4.7)\qquad P_s(D^0) := \int\int_{\{(y, \mathbf{z}) \in D^0\}} Q(y, d\mathbf{z}) \bar{F}(y) \mu^{-1}\, dy$$

(the integral of (2.4.4) with respect to the stationary distribution of $\{B_{m+n}\}$). Now use the decomposition of the renewal measure in Stone [**42**]: thus $U = U_1 + U_2$, where U_2 is a *finite* measure with finite $[l]$th order moment $\mu'_{[l]}$, and U_1 has a continuous density p satisfying $p(x) = 1/\mu + r(x)/\mu^2 + o(x^{-[l]})$ as $x \to \infty$, where $r(x) = \int_x^\infty (y - x) F(dy)$ for $x \geqslant 0$. Since $E(X^l) < \infty$,

$$r(x) \leqslant \int_x^\infty y F(dy) \leqslant x^{-(l-1)} \int_x^\infty y^l F(dy) = o(x^{-(l-1)}).$$

So $p(x) = 1/\mu + o(x^{-(l-1)})$. In our case U vanishes on $(-\infty, 0)$, so we may take $U_2(-\infty, 0) = 0$ and set

$$(2.4.8)\qquad p(x) = 1/\mu + q(x), \qquad x \in \mathbf{R},$$

where

$$(2.4.9)\qquad q(x) = -1/\mu \quad (x < 0), \qquad x^{l-1} q(x) \to 0 \quad (x \to \infty).$$

Fix $x \in [0, \frac{1}{2} n]$. From (2.4.6)–(2.4.8),

$$(2.4.10)\quad \begin{aligned} &P(D | F_m = x) - P_s(D^0) \\ &\quad = \int\int_{\{(y, \mathbf{z}) \in D^0\}} Q(y, d\mathbf{z}) \bar{F}(y) \{ U_2(n - x - dy) + q(n - x - y)\, dy \}. \end{aligned}$$

Write $\mu_l = E(X_1^l)$:

$$(2.4.11)\quad \begin{aligned} 0 &\leqslant \int\int_{\{(y, \mathbf{z}) \in D^0\}} Q(y, d\mathbf{z}) \bar{F}(y) U_2(n - x - dy) \leqslant \int_0^\infty \int_{\mathbf{R}^\infty} \\ &= \int_0^{n-x} \bar{F}(y) U_2(n - x - dy) \\ &\leqslant \int_0^{n/4} U_2(n - x - dy) + \bar{F}(n/4) \int_{n/4}^{n-x} U_2(n - x - dy) \\ &\leqslant U_2[n/4, \infty) + \bar{F}(n/4) U_2[0, \infty) \\ &\leqslant (4/n)^{[l]} \mu'_{[l]} + (4/n)^l \mu_l U_2[0, \infty). \end{aligned}$$

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

On the other hand, writing $q_0 := \sup_{t>0} |t^{l-1}q(t)|$, $q_1 := \sup_{t\in\mathbf{R}} |q(t)|$, both finite by (2.4.9) and continuity,

$$\begin{aligned}
&\left|\iint_{\{(y,\mathbf{z})\in D^0\}} Q(y, d\mathbf{z})\bar{F}(y)q(n-x-y)\,dy\right| \\
&\qquad \leqslant \int_0^\infty \bar{F}(y)|q(n-x-y)|\,dy \\
&\qquad \leqslant \int_0^{n/4} \bar{F}(y)(4/n)^{l-1}(n-x-y)^{l-1}|q(n-x-y)|\,dy \\
&\qquad\quad + \int_{n/4}^\infty \bar{F}(y)|q(n-x-y)|\,dy \qquad (2.4.12)\\
&\qquad \leqslant (4/n)^{l-1}q_0\int_0^{n/4} \bar{F}(y)\,dy + q_1\int_{n/4}^\infty \bar{F}(y)\,dy \\
&\qquad \leqslant (4/n)^{l-1}q_0\mu + q_1\int_{n/4}^\infty (x-n/4)F(dx) \\
&\qquad \leqslant (4/n)^{l-1}q_0\mu + (4/n)^{l-1}\mu_l q_1.
\end{aligned}$$

From (2.4.10)–(2.4.12), we thus have

$$|P(D|F_m = x) - P_s(D^0)| \leqslant \tfrac{1}{2}c'n^{-(l-1)} \tag{2.4.13}$$

for some constant c'. Now

$$P(B_{m+n} \in E) = \int_E \bar{F}(y)U(m+n-dy), \qquad E \in \mathfrak{B}, \tag{2.4.14}$$

so from (2.4.4), $P(D) = \iint_{\{(y,\mathbf{z})\in D^0\}} Q(y, d\mathbf{z})\bar{F}(y)U(m+n-dy)$, which is the same as the result of formally replacing $n - x$ by $m + n$ on the right of (2.4.6). The steps from (2.4.6) to (2.4.13) may be repeated for (2.4.14), whence

$$|P(D) - P_s(D^0)| \leqslant \tfrac{1}{2}c'n^{-(l-1)}.$$

Using (2.4.13), the proof of (2.4.2) is complete under LLT-I.

The lattice case is similar but simpler. Take x, y integers and replace dy in (2.4.7) by counting measure. For the renewal sequence $u_n = U(\{n\})$ $(n = 0, 1, \ldots)$, use $u_n = 1/\mu + o(n^{-(l-1)})$ (Stone [**41**], Karlin [**23**]) in place of the Stone decomposition above.

Proof of Theorem 4. Take cases LLT-I/II together. Denote by C_n the event $C \cap \{F_m \leqslant \frac{1}{2}n\}$, and let $C^0 \in \mathfrak{B}_{m+1}$ be the Borel set such that $C = \{\omega: (B_0(\omega), \ldots, B_m(\omega)) \in C^0\}$. Write the joint law of $(B_0, \ldots, B_m)$ as $R(d\mathbf{b})$, $\mathbf{b} = (b_0, \ldots, b_m)$. Given $B_m = b_m \geqslant 0$, the conditional distribution of F_m is

$$P(F_m \in A|B_m = b_m) = \int_A F(b_m + dx)/\bar{F}(b_m), \qquad A \in \mathfrak{B}, b_m \geqslant 0.$$

So

$$P(C_n)P(D) = P(D)\int_{C^0}\int_0^{n/2} \{F(b_m + dx)/\bar{F}(b_m)\}R(d\mathbf{b}),$$

$$P(C_n \cap D) = \int_{C^0}\int_0^{n/2} P(D|F_m = x)\{F(b_m + dx)/\bar{F}(b_m)\}R(d\mathbf{b}).$$

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

Subtract and use the lemma:

$$|P(C_n \cap D) - P(C_n)P(D)| \leqslant cn^{-(l-1)} \int_{C^0} \int_0^{n/2} \{F(b_m + dx)/\bar{F}(b_m)\} R(d\mathbf{b})$$
$$\leqslant cn^{-(l-1)} P(C_n) \leqslant cn^{-(l-1)}.$$

To remove the truncation of F_m, use

$$P(C \cap D) = P(C_n \cap D) + P\big(C \cap \{F_m > \tfrac{1}{2}n\} \cap D\big),$$
$$P(C) = P(C_n) + P\big(C \cap \{F_m > \tfrac{1}{2}n\}\big),$$

obtaining

$$\begin{aligned} &|P(C \cap D) - P(C)P(D)| \\ &\qquad \leqslant |P(C_n \cap D) - P(C_n)P(D)| + 2P\big(F_m > \tfrac{1}{2}n\big) \\ &\qquad \leqslant cn^{-(l-1)} + 2P\big(F_m > \tfrac{1}{2}n\big). \end{aligned} \tag{2.4.15}$$

To estimate the last term, we use $P(F_m > \frac{1}{2}n) = \int_0^m F(\frac{1}{2}n + m - x)U(dx)$. In case LLT-I, this gives by Stone's decomposition (setting $p_0 := \sup_{t \geqslant 0} p(t) < \infty$)

$$\begin{aligned} P\big(F_m > \tfrac{1}{2}n\big) &= \int_0^m \bar{F}\big(\tfrac{1}{2}n + m - x\big)U_2(dx) + \int_0^m \bar{F}\big(\tfrac{1}{2}n + m - x\big)p(x)\,dx \\ &\leqslant \bar{F}\big(\tfrac{1}{2}n\big)U_2[0, \infty) + p_0 \int_{n/2}^{\infty} \bar{F}(x)\,dx \\ &\leqslant \mu_l(2/n)^l U_2[0, \infty) + p_0(2/n)^{l-1}\mu_l \leqslant c_1 n^{-(l-1)}, \end{aligned}$$

where c_1 does not depend on m or n; similarly in the lattice case. Combining with (2.4.15), the conclusion of Theorem 4 follows.

2.5. *A Lai strong law for uniformly bounded strong mixing random variables.* The purpose of this subsection is to prove r-quick convergence to zero of $(\eta_1 + \cdots + \eta_n)/n^\alpha$, where the η_k are uniformly bounded strong-mixing random variables. The proof is a modification of that of [**27**, Theorem 2], in which we drop Lai's stationarity assumption but strengthen his moment assumptions to uniform boundedness. At the expense of some complication, we could have weakened uniform boundedness to a suitable uniform integrability condition.

Note that if $\{U_1, U_2, \ldots\}$ is any strong-mixing sequence and g_k: $\mathbf{R} \to \mathbf{R}$ are uniformly bounded measurable functions then $\{g_1(U_1), g_2(U_2), \ldots\}$ is a uniformly bounded strong-mixing sequence with the same (or better) mixing coefficients (for $\sigma(g_1(U_1), \ldots, g_k(U_k)) \subset \sigma(U_1, \ldots, U_k)$, etc.).

THEOREM 5. *Let $\eta_1, \eta_2, \ldots$ be strong mixing with mixing coefficients $\rho_n = O(n^{-\theta})$ for some $\theta > 1$, and assume $|\eta_k| \leqslant M$ for all k, where M is constant. Let α, p be any real numbers satisfying $\alpha > \frac{1}{2}$, $\theta > p > 1/\alpha$. If $\alpha \leqslant 1$ assume $E\eta_k = 0$ for all k. Then*

$$(\eta_1 + \cdots + \eta_n)/n^\alpha \to 0, \qquad (p\alpha - 1)\text{-}quickly.$$

PROOF. Set $V_0 = 0$, $V_n = \eta_1 + \cdots + \eta_n$, $n \geqslant 1$. When $\alpha > 1$ we have $|EV_n/n^\alpha| \leqslant Mn^{1-\alpha} \to 0$; hence the result of the theorem is equivalent to the same statement for r.v.s centred at expectations. So we centre at expectations, and henceforth

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

assume $E\eta_k = 0$ for all k. By [27, Lemma 4] it suffices to prove

$$\sum_{n=1}^{\infty} n^{p\alpha-2} P\Big(\max_{j=1,\dots,n} |V_j| \geqslant \varepsilon n^{\alpha}\Big) < \infty \quad \text{for all } \varepsilon > 0. \tag{2.5.1}$$

Choose δ satisfying $0 < \delta < 1 - p/\theta$; then since $\rho_n = O(n^{-\theta})$ we have

$$\sum_{n=1}^{\infty} n^{p\alpha-1} \rho_{n'} < \infty, \tag{2.5.2}$$

where $n' := [n^{\alpha(1-\delta)}]$. Take a positive integer k such that

$$k(2\alpha - 1) > p\alpha - 1. \tag{2.5.3}$$

Fix $\varepsilon > 0$. We denote $V_{ij} := \eta_{i+1} + \cdots + \eta_{i+j}$. Define

$$c_n(\varepsilon) = \sup_{r=0,1,\dots} P\Big(\max_{j=1,\dots,n} |V_{r,j}| > \varepsilon n^{\alpha}\Big) \leqslant 1.$$

For fixed n, r define $\tau_r = \inf\{j \geqslant 1\colon |V_{r,j}| \geqslant \frac{1}{2}\varepsilon n^{\alpha}/k\}$. Pick integers ν, i satisfying $0 \leqslant \nu \leqslant k-2$ and $1 \leqslant i \leqslant n$. Assuming n is large enough for $n^{\alpha\delta} \geqslant 2kM/\varepsilon$, we know that

$$\max_{j=1,\dots,n'} |V_{r+i-1,j}| \leqslant \tfrac{1}{2}\varepsilon n^{\alpha}/k. \tag{2.5.4}$$

If $\tau_r = i$ then

$$\max_{j=1,\dots,i-1} |V_{r,j}| \leqslant \tfrac{1}{2}\varepsilon n^{\alpha}/k. \tag{2.5.5}$$

If also it is the case that

$$\max_{j=1,\dots,n} |V_{r+i-1+n',j}| < \varepsilon n^{\alpha}(1 - (\nu+1)/k) \tag{2.5.6}$$

then it follows from (2.5.4)–(2.5.6), by addition, that

$$\max_{j=1,\dots,i-1+n'+n} |V_{r,j}| < \varepsilon n^{\alpha}(1 - \nu/k). \tag{2.5.7}$$

Hence if $\tau_r = i$, the negation of (2.5.7) implies the negation of (2.5.6), and in particular

$$\begin{aligned}\Big\{\tau_r = i, \max_{j=1,\dots,n} |V_{r,j}| \geqslant \varepsilon n^{\alpha}(1 - \nu/k)\Big\} \\ \subset \Big\{\tau_r = i, \max_{j=1,\dots,n} |V_{r+i-1+n',j}| \geqslant \varepsilon n^{\alpha}(1 - (\nu+1)/k)\Big\}.\end{aligned}$$

Therefore

$$\begin{aligned}&P\Big(\max_{j=1,\dots,n} |V_{r,j}| \geqslant \varepsilon n^{\alpha}(1 - \nu/k)\Big) \\ &\quad \leqslant \sum_{i=1}^{n} P\Big(\tau_r = i, \max_{j=1,\dots,n} |V_{r+i-1+n',j}| \geqslant \varepsilon n^{\alpha}(1 - (\nu+1)/k)\Big).\end{aligned} \tag{2.5.8}$$

By strong mixing, the latter sum is at most

$$\begin{aligned}\sum_{i=1}^{n} P(\tau_r = i) P\Big(\max_{j=1,\dots,n} |V_{r+i-1+n',j}| \geqslant \varepsilon n^{\alpha}(1 - (\nu+1)/k)\Big) + n\rho_{n'} \\ \leqslant P(\tau_r \leqslant n) c_n(\varepsilon(1 - (\nu+1)/k)) + n\rho_{n'} \\ \leqslant c_n\big(\tfrac{1}{2}\varepsilon/k\big) \cdot c_n(\varepsilon(1 - (\nu+1)/k)) + n\rho_{n'}.\end{aligned}$$

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

Substituting into (2.5.8) and taking the supremum over r gives

$$c_n(\varepsilon(1-\nu/k)) \leqslant c_n\left(\tfrac{1}{2}\varepsilon/k\right)\cdot c_n(\varepsilon(1-(\nu+1)/k)) + n\rho_{n'}.$$

This holds for $\nu = 0, 1, \ldots, k-2$, whence

$$c_n(\varepsilon) \leqslant \left\{c_n\left(\tfrac{1}{2}\varepsilon/k\right)\right\}^{k-1}\cdot c_n(\varepsilon/k) + n(k-1)\rho_{n'}.$$

Clearly $c_n(\varepsilon)$ is nonincreasing in ε, so we conclude

$$c_n(\varepsilon) \leqslant \left\{c_n\left(\tfrac{1}{2}\varepsilon/k\right)\right\}^{k} + nk\rho_{n'}. \tag{2.5.9}$$

Now by Ibragimov [**19**, Lemma 1.2], $|E(\eta_i\eta_{i+j})| \leqslant 4M^2\rho_j$, $i = 1, 2, \ldots,$ $j = 0, 1, \ldots$ (where $\rho_0 = 1$), whence by Serfling [**38**, Corollary A2],

$$E\left\{\left(\max_{j=1,\ldots,n} |V_{r,j}|\right)^2\right\} \leqslant 8M^2n(\log_2(2n))^2\sum_0^\infty \rho_i.$$

Chebychev's inequality gives

$$P\left(\max_{j=1,\ldots,n} |V_{r,j}| \geqslant \tfrac{1}{2}\varepsilon n^\alpha/k\right) \leqslant Kn^{1-2\alpha}(\log_2(2n))^2$$

for $K = 32(kM/\varepsilon)^2\sum_i \rho_i$, and so $c_n(\frac{1}{2}\varepsilon/k) \leqslant Kn^{1-2\alpha}(\log_2(2n))^2$. From (2.5.9),

$$c_n(\varepsilon) \leqslant K^k n^{-(2\alpha-1)k}(\log_2(2n))^{2k} + nk\rho_{n'},$$

and so $\sum_1^\infty n^{p\alpha-2}c_n(\varepsilon) < \infty$ on applying (2.5.2) and (2.5.3). From this the definition of $c_n(\varepsilon)$ yields (2.5.1).

2.6. *The case* $F(0+) = 0$. This subsection is devoted to proving Theorem 1 under the extra assumption $F(0+) = 0$. Under that assumption, λ in the statement of the theorem can be taken as $+\infty$, so r has to satisfy only $0 < r < \alpha(l-1) - 1$. Further, in the nonlattice case it will be evident that condition LLT-I suffices, rather than that $\mathcal{P}$ itself should have to have an absolutely continuous component.

Define N_t, B_t, F_t as in §§2.1, 2.4, and again write $\mathcal{F}_t := \sigma(N_u\colon 0 \leqslant u \leqslant t)$. (Actually we need $\mathcal{F}_n$ only for integer n. As $N_0 = 0$, $\mathcal{F}_0 = \{\varnothing, \Omega\}$.) Defining Y_n as in §2.1, we see that $\{Y_n, \mathcal{F}_n\}$ is an adapted sequence, so that $\{Y_n - E(Y_n|\mathcal{F}_{n-1}), \mathcal{F}_n\}$ is a martingale difference sequence. Further, the conditional distribution of F_{n-1} given $\mathcal{F}_{n-1}$ involves only B_{n-1}:

$$P(F_{n-1} > x|\mathcal{F}_{n-1}) = \bar{F}(B_{n-1}+x)/\bar{F}(B_{n-1}) \quad \text{a.s., } x \geqslant 0,$$

(recall $\bar{F} := 1 - F$), from which it follows that $E(Y_n|\mathcal{F}_{n-1}) = g_n(B_{n-1})$ a.s., $n = 1, 2, \ldots,$ for certain measurable functions, $g_1, g_2, \ldots$. The first lemma concerns these functions. Recall $U := \sum_0^\infty F^{(n)}$, $\sup|f(\cdot)| \leqslant M$.

LEMMA 2.6.1. $|g_n(b)| \leqslant MU(1) < \infty$ *for all* $b \geqslant 0$, $n = 1, 2, \ldots$.

PROOF.

$$\begin{aligned}|g_n(b)| &= \left|E\left(\sum_{N_{n-1}<k\leqslant N_n} f(S_k)|B_{n-1} = b\right)\right| \\ &\leqslant ME(N_n - N_{n-1}|B_{n-1} = b) \\ &= ME(E(N_n - N_{n-1}|F_{n-1})|B_{n-1} = b).\end{aligned}$$

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

Conditional on F_{n-1}, $N_n - N_{n-1}$ is at most N_1', the number of renewals in the time interval $[n - 1 + F_{n-1}, n + F_{n-1}]$. Since N_1' has the same distribution as $N_1 + 1$, with mean $U(1)$, the lemma follows.

LEMMA 2.6.2. *Fix r, α satisfying $\alpha > \frac{1}{2}$, $0 < r < \alpha(l - 1) - 1$. Then*

$$n^{-\alpha}\sum_1^n (Y_k - EY_k) \to 0, \quad r\text{-quickly}. \tag{2.6.1}$$

PROOF. Pick $\nu > \max(2, r/(\alpha - \frac{1}{2}))$. As in the previous proof, $|Y_n| \leqslant MN_1'$; hence $E(|Y_n|^\nu) \leqslant M^\nu E\{(N_1 + 1)^\nu\}$. It is well known that N_1 has all moments finite (cf. e.g. [**35**, p. 155]), so we have $\sup_n E(|Y_n|^\nu) < \infty$. Using Lemma 2.6.1 and Minkowski's inequality it follows that

$$\sup_n E(|Y_n - g_n(B_{n-1})|^\nu) < \infty.$$

Further, $Y_n - g_n(B_{n-1})$ is a martingale difference sequence; hence by Lemma 2.2.3,

$$n^{-\alpha}\sum_1^n \{Y_k - g_k(B_{k-1})\} \to 0, \quad r\text{-quickly}. \tag{2.6.2}$$

On the other hand, consider the sequence $g_k(B_{k-1}) - Eg_k(B_{k-1})$, $k = 1, 2, \ldots,$ which is uniformly bounded by $2MU(1)$, and centred at expectations. By Theorem 4, $\{B_{k-1}\}_1^\infty$ is strong mixing with mixing coefficients $\rho_n = O(n^{-(l-1)})$; hence so is the sequence $g_k(B_{k-1}) - Eg_k(B_{k-1})$. By Theorem 5, since $r = p\alpha - 1$ where $l - 1 > p > 1/\alpha$,

$$n^{-\alpha}\sum_1^n \{g_k(B_{k-1}) - Eg_k(B_{k-1})\} \to 0, \quad r\text{-quickly}. \tag{2.6.3}$$

However $Eg_k(B_{k-1}) = E(E(Y_k|\mathcal{F}_{k-1})) = EY_k$, and on inserting this into (2.6.3), and combining with (2.6.2) using Lemma 2.2.1, the result follows.

PROOF OF THEOREM 1 (WHEN $F(0+) = 0$). We can write (2.6.1) as

$$n^{-\alpha}\left\{\sum_1^{N(n)} f(S_k) - E\sum_1^{N(n)} f(S_k)\right\} \to 0, \quad r\text{-quickly}, \tag{2.6.4}$$

recalling that $N(n)$ is an alternative notation for N_n. We prove

$$n^{-\alpha}\left\{\sum_1^{N(n)} f(S_k) - \sum_1^{[n/\mu]} f(S_k)\right\} \to 0, \quad r\text{-quickly}, \tag{2.6.5}$$

and

$$n^{-\alpha}E\left\{\sum_1^{N(n)} f(S_k) - \sum_1^{[n/\mu]} f(S_k)\right\} \to 0. \tag{2.6.6}$$

The convergence in (2.6.6) may be considered as r-quick convergence of degenerate r.v.s; hence (2.6.4)–(2.6.6) together imply (1.1), using Lemma 2.2.1.

To prove (2.6.5) note that

$$n^{-\alpha}\left|\sum_1^{N(n)} f(S_k) - \sum_1^{[n/\mu]} f(S_k)\right| \leqslant Mn^{-\alpha}(|N_n - n/\mu| + 1)$$

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

and the latter tends to zero ($l\alpha - 1$)-quickly for $\frac{1}{2} < \alpha \leqslant 1$, ∞-quickly for $\alpha > 1$, so r-quickly in either case since $r < l\alpha - 1$.

For (2.6.6), since $\sigma^2 = \operatorname{var} X_1 < \infty$ we have $EN_t = U(t) - 1 = t/\mu + O(1)$ [**16**, **I**, XIII (12.2) and **II**, XI (3.1)], and also $\operatorname{var} N_t \sim \sigma^2 t/\mu^3$ [**16**, **II**, XI.10, Example 13]. Thus

$$\begin{aligned}\left| n^{-\alpha} E\left\{ \sum_1^{N(n)} f(S_k) - \sum_1^{[n/\mu]} f(S_k) \right\} \right| &\leqslant Mn^{-\alpha}(E|N_n - n/\mu| + 1) \\ &\leqslant Mn^{-\alpha}(E|N_n - EN_n| + |EN_n - n/\mu| + 1) \\ &\leqslant Mn^{-\alpha}\left\{(\operatorname{var} N_n)^{1/2} + O(1)\right\} \\ &= Mn^{-\alpha}\left\{(\sigma^2 n/\mu^3)^{1/2}(1 + o(1)) + O(1)\right\} \to 0\end{aligned}$$

since $\alpha > \frac{1}{2}$. The result follows.

2.7. *Ladder-height distributions*. We return to the framework of §1, with no special assumptions. Set $L_1 := \min\{k: S_k > 0\}$ and define $A := \sum_1^\infty n^{-1}P(S_n \leqslant 0)$. Then L_1 is proper and has finite expectation if and only if $A < \infty$ [**16**, **II**, p. 416], and in that case $EL_1 = e^A$. A sufficient condition is that $0 < EX_1 < \infty$ [**16**, **II**, p. 397].

THEOREM 6. *Assume $A < \infty$. If F has an absolutely continuous component then so does G, the d.f. of $S(L_1) = S_{L_1}$.*

PROOF. Denote the Borel sets on $(0, \infty)$ by $\mathfrak{B}^+$. Denote the measures on $(\mathbf{R}, \mathfrak{B})$ induced by F and its absolutely continuous component by ν, ν_c respectively. If the support of ν_c intersects $(0, \infty)$ then immediately

$$\forall B \in \mathfrak{B}^+, \qquad P(S(L_1) \in B) \geqslant P(X_1 \in B) \geqslant \nu_c(B),$$

so G has a nontrivial absolutely continuous component. We therefore assume from now on that the support of ν_c is contained in $(-\infty, 0]$. Absolute continuity means that

$$\nu_c(B) = \int_B w(x)\,dx, \quad \text{all } B \in \mathfrak{B}, \tag{2.7.1}$$

for some nonnegative measurable function w. By assumption ν_c has a positive total mass, so there is a set of positive measure on which $w > 0$. By modifying w on a set of zero measure if need be, we may take it that there exists an open interval $(-b, -a)$, with $0 \leqslant a < b$, in which $w > 0$. The modification preserves (2.7.1).

Since $A < \infty$, F has some point of increase $x_0 > 0$, that is, $\nu(x_0 - \varepsilon, x_0 + \varepsilon) > 0$ for all $\varepsilon > 0$. We may find positive integers j, k such that the point $x_1 = (2j + 1)x_0/(2k)$ is contained in (a, b). So $-b < -x_1 < -a$ and $kx_1 = (j + \frac{1}{2})x_0$. Set $\varepsilon = \frac{1}{2}x_0/(k + j + 1)$. Let ν_{cr} denote the restriction of ν_c to the interval $(-x_1 - \varepsilon, -x_1 + \varepsilon)$ and let ν_r denote the restriction of ν to the interval $(x_0 - \varepsilon, x_0 + \varepsilon)$. Both ν_{cr} and ν_r have their total masses positive. Denote by E the event that $X_i \in (-x_1 - \varepsilon, -x_1 + \varepsilon)$ for $i = 1, \ldots, k$ and $X_i \in (x_0 - \varepsilon, x_0 + \varepsilon)$ for $i = k + 1, \ldots, k + j + 1$. On E the path of the sequence $\{S_n\}_0^\infty$ has downward jumps at

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

$n = 1, \ldots, k$, then upward jumps at $n = k + 1, \ldots, k + j + 1$; further,

$$S_{k+j} < k(-x_1 + \varepsilon) + j(x_0 + \varepsilon) = (k + j)\varepsilon - \tfrac{1}{2}x_0 < 0,$$

$$S_{k+j+1} > k(-x_1 - \varepsilon) + (j + 1)(x_0 - \varepsilon) = -(k + j + 1)\varepsilon + \tfrac{1}{2}x_0 > 0.$$

Thus on the event E, $S(L_1) = S_{k+j+1}$. So for every $B \in \mathfrak{B}^+$,

$$P(S(L_1) \in B) \geqslant P(E \cap \{S_{k+j+1} \in B\}) \geqslant (\nu_{cr}^{(k)} * \nu_r^{(j+1)})(B). \tag{2.7.2}$$

The convolution $\nu_{cr}^{(k)} * \nu_r^{(j+1)}$ has support contained in $(0, \infty)$, has positive total mass, and is absolutely continuous since ν_{cr} is. By (2.7.2) the distribution of $S(L_1)$ has a nontrivial absolutely continuous component, as claimed.

Note. We need this theorem only to ensure that the distribution of first ladder-height satisfies LLT-I. Whether this weaker conclusion follows from the weaker condition that $\mathcal{P}$ satisfies LLT-I is an open question.

2.8. *Proof of Theorem* 1. Fix r and α satisfying the inequalities specified in the theorem. Define strict ladder epochs

$$L_0 = 0, \quad L_n = \min\{k: k > L_{n-1}, S(k) > S(L_{n-1})\}, \qquad n = 1, 2, \ldots.$$

These are a.s. finite stopping times. Define the interladder times as $M_k = L_k - L_{k-1}$, $k = 1, 2, \ldots$. We shall use the fact (cf. [**10**, p. 136]) that the random vectors

$$(M_k, X(L_{k-1} + 1), X(L_{k-1} + 2), \ldots, X(L_k)), \qquad k = 1, 2, \ldots, \tag{2.8.1}$$

are i.i.d. Setting $S_k' = S(L_k)$, $k = 0, 1, \ldots$, $X_k' = S_k' - S_{k-1}' = \sum_{L_{k-1} < j \leqslant L_k} X_j$, it follows that the random vectors (M_k, X_k'), $k = 1, 2, \ldots$, are i.i.d. Since $M_1 = L_1$, $X_1' = S(L_1)$, a result of Gut [**17**, Theorem 2.1] gives $E(M_1^\lambda) < \infty$, $E((X_1')^l) < \infty$. Gut proves also the converse assertion, which shows in the present context that our moment assumptions on X_1^+, X_1^- are necessary for these conclusions. We define $R_0 = 0$, $R_k = \sum_{L_{k-1} < j \leqslant L_k} f(S_j)$, $k = 1, 2, \ldots$.

LEMMA 2.8.1. *$n^{-\alpha}\sum_1^n (R_k - ER_k) \to 0$, r-quickly.*

PROOF. Define the pre-L_n σ-algebra by

$$\mathcal{G}_n = \{B \in \mathcal{F}: B \cap \{L_n = k\} \in \sigma(X_1, \ldots, X_k), k = 0, 1, \ldots\}.$$

Then $\{R_k, \mathcal{G}_k\}_0^\infty$ is an adapted sequence and $R_k - E(R_k|\mathcal{G}_{k-1})$ is a martingale difference sequence. The conditional expectation may be written

$$E(R_k|\mathcal{G}_{k-1}) = E\left(\sum_{j=1}^{M_k} f(S(L_{k-1}) + X(L_{k-1} + 1) + \cdots + X(L_{k-1} + j))|\mathcal{G}_{k-1}\right)$$

and by (2.8.1) this evaluates as $g(S(L_{k-1}))$ where the function g is defined by $g(s) = E\sum_{j=1}^{M_1} f(s + S_j)$, $s \in \mathbf{R}$. Note that $|g(s)| \leqslant MEM_1 = Me^A < \infty$, where $A = \sum_1^\infty n^{-1}P(S_n \leqslant 0)$ as in §2.7. So g is a bounded measurable function on $\mathbf{R}$.

Next, $|R_k| \leqslant MM_k$, so

$$E(|R_k|^\lambda) \leqslant M^\lambda E(M_k^\lambda) = M^\lambda E(M_1^\lambda) < \infty.$$

The martingale difference sequence $R_k - g(S_{k-1}')$ therefore satisfies the conditions of Lemma 2.2.3 with $\nu = \lambda$, whence

$$n^{-\alpha}\sum_1^n \{R_k - g(S_{k-1}')\} \to 0, \quad r\text{-quickly}. \tag{2.8.2}$$

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

Now $S_k' = \Sigma_1^k X_j'$ where the X_j' are i.i.d. with $E((X_1')^l) < \infty$, and the d.f. G of X_1' satisfies $G(0+) = 0$. If F is lattice then obviously so is G, while if F has an absolutely continuous component then so does G by Theorem 6. Thus the result of §2.6—the special-case proof of Theorem 1—applies, and since $r < \alpha(l-1) - 1$ we conclude

$$n^{-\alpha}\sum_1^n \{g(S_k') - Eg(S_k')\} \to 0, \quad r\text{-quickly}.$$

So

$$n^{-\alpha}\sum_1^n \{g(S_{k-1}') - Eg(S_{k-1}')\}$$

$$= \left(\frac{n-1}{n}\right)^\alpha \cdot (n-1)^{-\alpha} \cdot \sum_1^{n-1} \{g(S_k') - Eg(S_k')\} \tag{2.8.3}$$

$$\to 0, \quad r\text{-quickly}.$$

However, $Eg(S_{k-1}') = E(E(R_k|\mathcal{G}_{k-1})) = ER_k$. Substituting into (2.8.3), then combining with (2.8.2) using Lemma 2.2.1, the present lemma follows.

PROOF OF THEOREM 1. The conclusion of Lemma 2.8.1 may be written

$$n^{-\alpha}\left\{\sum_1^{L_n} f(S_k) - E\sum_1^{L_n} f(S_k)\right\} \to 0, \quad r\text{-quickly}. \tag{2.8.4}$$

We prove

$$n^{-\alpha}\left\{\sum_1^{L_n} f(S_k) - \sum_1^{[ne^A]} f(S_k)\right\} \to 0, \quad r\text{-quickly}, \tag{2.8.5}$$

and

$$n^{-\alpha}E\left\{\sum_1^{L_n} f(S_k) - \sum_1^{[ne^A]} f(S_k)\right\} \to 0. \tag{2.8.6}$$

Note the similarity to (2.6.4)–(2.6.6). Just as in that proof, (2.8.4)–(2.8.6) may be combined, using Lemma 2.2.1, to give

$$n^{-\alpha}\sum_1^{[ne^A]} \{f(S_k) - Ef(S_k)\} \to 0, \quad r\text{-quickly},$$

and it is straightforward—using the uniform boundedness of the summands—to deduce (1.1).

To prove (2.8.5), note that the absolute value of its left-hand side is at most $n^{-\alpha}M(1 + |L_n - ne^A|)$. Since $L_n = M_1 + \cdots + M_n$ where the M_k are i.i.d. with $EM_1 = e^A$, $E(M_1^\lambda) < \infty$, the Lai strong law [**26**, (6.5)] yields $n^{-\alpha}(L_n - ne^A) \to 0$ $(\lambda\alpha - 1)$-quickly. But $r < \lambda\alpha - \frac{1}{2}\lambda \leqslant \lambda\alpha - 1$, so the same is true r-quickly, which gives (2.8.5). To prove (2.8.6) the same second-moment argument as was used for (2.6.6) is applicable. Theorem 1 is proved.

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

3. Proofs of Theorems 2 and 2′. Define $S_t = S(t) := S_{[t]}$ for $t \geqslant 0$. Let $W(t) = W_t$ be a Wiener process with mean μ and variance σ^2. For $T > 0$ define

$$\lambda_T := T^{-1}\int_0^T f(\mu t)\,dt, \qquad \mu_T := T^{-1}\int_0^T f(S_t)\,dt, \qquad \nu_T := T^{-1}\int_0^T f(W_t)\,dt.$$

We prove Theorem 2 by showing $E(\mu_T - \nu_T) = o(T^{-\delta/2+\varepsilon})$ and $E\nu_T - \lambda_T = o(T^{-1/2+\varepsilon})$, $T \to \infty$. The notation, and the idea of using ν_T, come from papers of Davydov (e.g. **[12]**, **[13]**). As does Davydov, we proceed by proving equiconvergence of the variation-norm distance between relevant distributions. Our work differs from his in that we need the *rate* of convergence rather than just convergence itself; however only for univariate distributions.

For a signed measure ν on the Borel algebra $\mathfrak{B}$ in $\mathbf{R}$, define the variation norm by $\|\nu\| := \frac{1}{2}\int_{\mathbf{R}} |\nu|\,(dx)$. The following properties will be needed.

(a) For probability measures ν_1, ν_2, part of the result known as Scheffé's lemma gives

$$\|\nu_1 - \nu_2\| = \frac{1}{2}\int_{\mathbf{R}} |\nu_1(dx) - \nu_2(dx)| = \sup\{\nu_1(B) - \nu_2(B) \colon B \in \mathfrak{B}\}.$$

The supremum is attained on some $A^+ \in \mathfrak{B}$. If ν_1, ν_2 have densities h_1, h_2 then one may identify A^+ with (for instance) $\{x \colon h_1(x) > h_2(x)\}$.

(b) Variation norm is preserved under isomorphism of the measure space $(\mathbf{R}, \mathfrak{B}, \nu)$.

(c) Let $\Phi^{a,b}$ denote the d.f., and $\phi^{a,b}$ the density, of the normal distribution of mean a and variance b, and write $\Phi := \Phi^{0,1}$, $\phi := \phi^{0,1}$. Symbols for d.f.s are also to denote the corresponding Lebesgue-Stieltjes measures. Then

$$\|\Phi^{\varepsilon,1} - \Phi\| = O(\varepsilon), \quad \varepsilon\downarrow 0; \qquad \|\Phi^{0,1-\varepsilon} - \Phi\| = O(\varepsilon^{1/2}), \quad \varepsilon\downarrow 0.$$

(Evaluate the variation norms by integrating the difference of the densities over the region where that of (say) Φ is the greater.)

Write F_t, Φ_t for the d.f.s of S_t, W_t respectively. Proposition 3.1 below establishes the variation-norm closeness of F_t and Φ_t, and Proposition 3.2 deduces that $E\mu_T$ and $E\nu_T$ are close. The assumption $\mu > 0$ is superfluous for these results and will not be used. In the lattice case it is convenient (cf. Davydov **[12**, p. 440**]**) to work with a random variable $S_t + U$ instead of with S_t, where U is uniformly distributed on [0, 1], independent of S_t. Because of our standing assumption in the lattice case that f is constant on each interval $[k, k+1)$, we have $f(S_t) = f(S_t + U)$ a.s. Let G_t denote the d.f. of $S_t + U$.

PROPOSITION 3.1. *Assume* $E(|X_1|^{2+\delta}) < \infty$ *for some* δ *with* $0 < \delta \leqslant 1$. *If* $\mathcal{P}$ *satisfies* LLT-I *then*

$$\|F_t - \Phi_t\| = O(t^{-\delta/2}), \qquad t \to \infty, \tag{3.1}$$

while if $\mathcal{P}$ *satisfies* LLT-II *then, for every* θ *satisfying* $0 < \theta < \delta$,

$$\|G_t - \Phi_t\| = O(t^{-\theta/2}), \qquad t \to \infty. \tag{3.2}$$

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

PROOF. Assume LLT-I. We may write $F_n = H_n + J_n$ where H_n, J_n are improper d.f.s (nondecreasing functions that tend to 0 at $-\infty$) with H_n absolutely continuous and J_n singular. By LLT-I, $J_m(\infty) < 1$ for some m. Write $b := (J_m(\infty))^{1/m} < 1$. For $n \geqslant 1$ write $n = km + r$, $0 \leqslant r < m$; then $F_n = (F_m)^{(k)} * F_r = (H_m + J_m)^{(k)} * F_r$, so that J_n, the singular component of F_n, is a component of $J_m^{(k)} * F_r$, whence

$$J_n(\infty) \leqslant (J_m(\infty))^k \leqslant (J_m(\infty))^{(n/m)-1} \leqslant b^{n-m} = O(n^{-\delta/2}).$$

But $\|F_n - \Phi_n\| = \|F_n - \Phi^{n\mu,n\sigma^2}\| \leqslant \|H_n - \Phi^{n\mu,n\sigma^2}\| + J_n(\infty)$, and $\|H_n - \Phi^{n\mu,n\sigma^2}\| = O(n^{-\delta/2})$ by Sirazhdinov and Mamatov [**39**, Theorem 2]. Thus $\|F_n - \Phi_n\| = O(n^{-\delta/2})$. Taking any $t > 0$ and writing $n = [t]$, the conclusion (3.1) follows on writing

$$\begin{aligned}\|F_t - \Phi_t\| &= \left\|F_n - \Phi^{t\mu,t\sigma^2}\right\| \\ &\leqslant \left\|F_n - \Phi^{n\mu,n\sigma^2}\right\| + \left\|\Phi^{n\mu,n\sigma^2} - \Phi^{n\mu,t\sigma^2}\right\| + \left\|\Phi^{n\mu,t\sigma^2} - \Phi^{t\mu,t\sigma^2}\right\|\end{aligned}$$

and applying (**b**) and (**c**) above to the last terms.

Turning to the lattice case, we shall show that

$$\left\|G_n - \Phi^{n\mu,n\sigma^2}\right\| = O(n^{-\theta/2}), \qquad n \to \infty, \tag{3.3}$$

whence the conclusion (3.2) follows exactly as in the last step of the proof of (3.1). The d.f. G_n is absolutely continuous and its density g_n may be taken as $g_n(x) = P(S_n = k)$, $k \leqslant x < k + 1$, $k \in \mathbf{Z}$. So (3.3) is equivalent to

$$\int_{-\infty}^{\infty} \left|g_n(x) - \phi^{n\mu,n\sigma^2}(x)\right| dx = O(n^{-\theta/2}). \tag{3.4}$$

By [**20**, Theorem 4.5.3], for $\delta < 1$,

$$\sup_{k \in \mathbf{Z}} \left|\sigma n^{1/2} P(S_n = k) - \phi((k - n\mu)/(\sigma n^{1/2}))\right| = O(n^{-\delta/2}),$$

and since $\phi(\cdot)$ has a bounded derivative we deduce

$$\sup_{x \in \mathbf{R}} \left|\sigma n^{1/2} g_n(x) - \phi((x - n\mu)/(\sigma n^{1/2}))\right| = O(n^{-\delta/2}),$$

that is, $\Delta_n := \sup_{x \in \mathbf{R}} \sigma n^{1/2} |g_n(x) - \phi^{n\mu,n\sigma^2}(x)| = O(n^{-\delta/2})$. This extends to the case $\delta = 1$, using Petrov [**32**, Theorem 16, p. 207]. The rest of the proof is an extension of that in [**20**, pp. 124–125]. Define $E := \{x: |x - n\mu| \leqslant \sigma n^{1/2} \Delta_n^{-1+\theta/\delta}\}$; then

$$\int_E \left|g_n(x) - \phi^{n\mu,n\sigma^2}(x)\right| dx \leqslant 2\Delta_n^{\theta/\delta} = O(n^{-\theta/2}). \tag{3.5}$$

On the other hand since $\Phi(-x) + 1 - \Phi(x) = o(\exp(-\frac{1}{2}x^2))$ we have

$$\begin{aligned}\int_{\mathbf{R}\setminus E} \phi^{n\mu,n\sigma^2}(x)\, dx &= \Phi\left(-\Delta_n^{-1+\theta/\delta}\right) + 1 - \Phi\left(\Delta_n^{-1+\theta/\delta}\right) \\ &= O(n^{-\theta/2}),\end{aligned} \tag{3.6}$$

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

whence

$$\begin{aligned}\int_{\mathbf{R}\setminus E} g_n(x)\,dx &= 1 - \int_E g_n(x)\,dx \\ &\leqslant 1 - \int_E \phi^{n\mu,n\sigma^2}(x)\,dx + \int_E \left|g_n(x) - \phi^{n\mu,n\sigma^2}(x)\right|\,dx \\ &= O(n^{-\theta/2}).\end{aligned}$$

Combining this with (3.5) and (3.6) we obtain (3.4), and hence (3.2).

REMARK. The above-mentioned theorem of Petrov may be used directly to prove (3.2) for the case $\delta = 1$, and indeed it gives a strengthened version in which $\theta = \delta$. The result is presumably true in general with $\theta = \delta$, but the weaker form given suffices here.

PROPOSITION 3.2. *Under the conditions of Theorem* 2, *for every* $\varepsilon > 0$,

$$E(\mu_T - \nu_T) = o(T^{-\delta/2+\varepsilon}), \qquad T \to \infty. \tag{3.7}$$

PROOF. Take ε with $0 < \varepsilon < \delta$ and set $\theta = \delta - \varepsilon$; then by Proposition 3.1 we may choose a constant C large enough so that under LLT-I, $\|F_t - \Phi_t\| \leqslant Ct^{-\theta/2}$, all $t > 0$, while under LLT-II, $\|G_t - \Phi_t\| \leqslant Ct^{-\varepsilon/2}$, all $t > 0$. Now under LLT-I,

$$\begin{aligned}|E(\mu_T - \nu_T)| &= \left|\int_0^1 E\{f(S_{Tu}) - f(W_{Tu})\}\,du\right| \\ &= \left|\int_0^1 \int_{-\infty}^{\infty} f(x)\{P(S_{Tu} \in dx) - P(W_{Tu} \in dx)\}\,du\right| \\ &\leqslant M\int_0^1 \int_{-\infty}^{\infty} |P(S_{Tu} \in dx) - P(W_{Tu} \in dx)|\,du \\ &= M\int_0^1 2\|F_{Tu} - \Phi_{Tu}\|\,du \\ &\leqslant 2M\int_0^1 C(Tu)^{-\theta/2}\,du = O(T^{-\theta/2}) = o(T^{-\delta/2+\varepsilon}).\end{aligned}$$

When $\mathcal{P}$ is lattice the same calculations are valid, with $S_{Tu} + U$, G_{Tu} replacing S_{Tu}, F_{Tu} respectively. This proves (3.7).

We turn now to the convergence to zero of $E\nu_T - \lambda_T$. Let ξ be uniformly distributed over [0, 1], independent of the process $\{W_t\}_{t\geqslant 0}$. Observe that λ_T may be considered as $Ef(\mu T\xi)$ whereas $E\nu_T$ equals $Ef(W(T\xi))$. This motivates the following result. In an earlier version of the paper we proved a multidimensional version (see §5.4 below).

PROPOSITION 3.3. *Assume* $\mu > 0$, $\sigma^2 < \infty$. *With* ξ *as above, let* $\mathcal{K}_T$, $\mathcal{L}_T$ *denote the distributions of* $\mu T\xi$, $W(T\xi)$ *respectively. Then for every* $\varepsilon > 0$,

$$\|\mathcal{L}_T - \mathcal{K}_T\| = o(T^{1/2+\varepsilon}), \qquad T \to \infty. \tag{3.8}$$

PROOF. Set $B_t := (W_t - \mu t)/\sigma$, so that $\{B_t\}_{t\geqslant 0}$ is standard Brownian motion, and $\mathcal{L}_T$ is the distribution of $\mu T\xi + \sigma B(T\xi)$. Applying a common scale transformation and using remark **(b)** we have

$$\|\mathcal{L}_T - \mathcal{K}_T\| = \|\mathcal{N}_T - \mathcal{M}\| \tag{3.9}$$

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

where $\mathfrak{M}$ is the uniform distribution on [0, 1] and $\mathfrak{N}_T$ is the distribution of $\xi + \sigma(\mu T)^{-1}B(T\xi)$. Using the scale property of Brownian motion, $\mathfrak{N}_T$ is also the distribution of $\xi + B(\xi/T')$, where $T' := \mu^2 T/\sigma^2$. Thus $\mathfrak{N}_T$ has probability density

$$(3.10)\quad s_T(x) := \int_0^1 (2\pi y/T')^{-1/2} \exp\left\{-\tfrac{1}{2}(x-y)^2 T'/y\right\} dy, \qquad x \in \mathbf{R}.$$

Fix $x > 0$. The integrand in (3.10), as a function of y, defines a probability density on $\mathbf{R}^+$, indeed one of the 'random walk' densities [**21**, p. 149]. So if the integral in (3.10) were extended to the interval $(0, \infty)$, its value would be 1; hence its actual value is less. Thus the density of $\mathfrak{M}$ is greater than that of $\mathfrak{N}_T$ on the interval (0, 1), and not elsewhere, whence the variation-norm distance evaluates as

$$\|\mathfrak{N}_T - \mathfrak{M}\| = \int_{\mathbf{R}\setminus(0,1)} s_T(x)\,dx = P(\xi + B(\xi/T') \notin (0, 1)).$$

Fix ε satisfying $0 < \varepsilon < 1$. Now

$$\begin{aligned} P(\xi + B(\xi/T') \notin (0,1)) &\leqslant P(\xi < T^{-1/2+\varepsilon/2}) + P(\xi > 1 - T^{-1/2+\varepsilon/2}) \\ &\quad + P(|B(\xi/T')| \geqslant T^{-1/2+\varepsilon/2}) \\ &\leqslant 2T^{-1/2+\varepsilon/2} + P(|B(1/T')| \geqslant T^{-1/2+\varepsilon/2}) \end{aligned}$$

(since $|B(1/T')|$ is stochastically larger than $|B(\xi/T')|$)

$$= 2T^{-1/2+\varepsilon/2} + \Phi(-(\mu/\sigma)T^{\varepsilon/2}) + 1 - \Phi((\mu/\sigma)T^{\varepsilon/2}).$$

Since $\Phi(-x) = 1 - \Phi(x)$ is asymptotically smaller than any power of x we conclude that the final right-hand side above is $o(T^{-1/2+\varepsilon})$, which with (3.9) and (3.10) gives the result.

PROOF OF THEOREM 2. Since $E\nu_T - \lambda_T = E\{f(W(T\xi)) - f(\mu T\xi)\}$ it is easy to show $E\nu_T - \lambda_T = o(T^{-1/2+\varepsilon})$ by the method used to prove (3.7). With (3.7) the conclusion follows.

PROOF OF THEOREM 2′. For $2 \leqslant l \leqslant 3$ we have $2 - \frac{1}{2}l \geqslant (l-1)^{-1}$. Hence any value of α satisfying the conditions of Theorem 2 will satisfy that in Theorem 1 also. Set $\delta = \min(l, \lambda, 3) - 2$; then the condition of Theorem 2 reduces to $\alpha > 1 - \frac{1}{2}\delta$. Further, $E(|X_1|^{2+\delta}) < \infty$, so Theorem 2 applies and we may take ε in Theorem 2 to be $\alpha - 1 + \frac{1}{2}\delta$. Then Theorem 2 becomes

$$n^{-\alpha}\left\{E\sum_1^n f(S_k) - \int_0^n f(\mu x)\,dx\right\} \to 0,$$

which with Theorem 1 yields (1.4) as required.

4. The Poisson case. Consider the special case where $\mathfrak{P}$ is an exponential distribution, of parameter 1 say. Then $\max\{k\colon S_k \leqslant t\}$ gives a Poisson process of rate 1.

We note first a simple self-contained proof of the equivalence of statements (I) and (IV) in §1 (cf. Corollary 1.1) for this case. First,

$$(4.1)\qquad n^{-1}\sum_1^n f(S_k) \to c \quad \text{a.s. } (n \to \infty)$$

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

is equivalent to

$$N_t^{-1} \sum_1^{N(t)} f(S_k) \to c \quad \text{a.s. } (t \to \infty), \tag{4.2}$$

as the successive values of the two left-hand sides coincide. Next,

$$Y_n := \sum_{N(n-1) < k \leqslant N(n)} f(S_k) = \sum_{n-1 < S_k \leqslant n} f(S_k)$$

is the sum of the f-values at the Poisson epochs in the time interval $(n-1, n]$. By independence of the Poisson point process over disjoint time intervals, the Y_k are *independent*. By the strong law, $N_t/t \to 1$ $(t \to \infty)$ a.s.; hence (4.1), (4.2) are equivalent to

$$n^{-1} \sum_1^n Y_k \to c \quad \text{a.s. } (n \to \infty). \tag{4.3}$$

Recall $M := \sup|f(\cdot)| < \infty$; then $|Y_n| \leqslant M(N_n - N_{n-1})$, $\operatorname{var} Y_n \leqslant M^2$, $\sum_1^\infty n^{-2} \operatorname{var} Y_n < \infty$. So by a result of Kolomogorov [**43**, p. 165],

$$n^{-1} \sum_1^n (Y_k - EY_k) \to 0 \quad \text{a.s. } (n \to \infty).$$

So (4.1)–(4.3) are equivalent to

$$n^{-1} \sum_1^n EY_k \to c \qquad (n \to \infty). \tag{4.4}$$

But, conditional on m Poisson points falling in $(n-1, n]$, these are independently uniformly distributed over $(n-1, n]$, so

$$EY_n = \sum_0^\infty (e^{-1}/m!) \cdot m \int_{n-1}^n f(y)\, dy = \int_{n-1}^n f.$$

Thus (4.4) is $f(\cdot) \to c$ (C_1), as required.

The same method yields a CLT for $f(S_n)$. We assume that $\int_0^n \{f(y)\}^2\, dy = \int_0^n f^2 \to \infty$ $(n \to \infty)$. By the method used to calculate EY_n, or by generating functions,

$$\operatorname{var} Y_k = \int_{k-1}^k f^2, \qquad E\left\{\left(Y_k - \int_{k-1}^k f\right)^3\right\} = \int_{k-1}^k f^3.$$

Now $|\int_0^n f^3| \leqslant M \int_0^n f^2 = o(\{\int_0^n f^2\}^{3/2})$ so we have Liapounov's condition for the Y_k. By Liapounov's theorem,

$$\left(\int_0^n f^2\right)^{-1/2} \left(\sum_1^n Y_k - \int_0^n f\right) \xrightarrow{d} \Phi \qquad (n \to \infty),$$

or

$$\left(\int_0^n f^2\right)^{-1/2} \left\{\sum_1^{N(n)} f(S_k) - \int_0^n f\right\} \xrightarrow{d} \Phi. \tag{4.5}$$

By the strong law, $S_n/n \to 1$ a.s., and $N(S_n) = n$. To replace n by S_n we use Anscombe's method for random time-changes, of which [**4**, Theorem 17.1] is a

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

formulation suitable for our case of independent nonidentically distributed summands. Thus

$$\left(\int_0^{S(n)} f^2\right)^{-1/2}\left\{\sum_1^n f(S_k) - \int_0^{S(n)} f\right\} \xrightarrow{d} \Phi.$$

If we assume further

$$\liminf_{n\to\infty} n^{-1}\int_0^n f^2 > 0 \tag{4.6}$$

then $|\int_0^{S(n)} f^2/\int_0^n f^2 - 1| \leqslant Mn^{-1}|S_n - n| \cdot n/\int_0^n f^2 \to 0$ a.s.; hence we may replace $\int_0^{S(n)} f^2$ by $\int_0^n f^2$ in the above to obtain

$$\left(\int_0^n f^2\right)^{-1/2}\left\{\sum_1^n f(S_k) - \int_0^{S(n)} f\right\} \xrightarrow{d} \Phi,$$

the CLT for $F(S_n)$ (note the random centring).

Again, assuming (4.6), Petrov's LIL [**43**, Corollary 5.2.3] yields

$$\limsup_{n\to\infty}\left\{2\int_0^n f^2 \cdot \log\log\int_0^n f^2\right\}^{-1/2}\sum_1^n (Y_k - EY_k) = 1 \quad \text{a.s.},$$

or

$$\limsup_{n\to\infty}\left\{2\int_0^n f^2 \log\log\int_0^n f^2\right\}^{-1/2}\left\{\sum_1^{N(n)} f(S_k) - \int_0^n f\right\} = 1 \quad \text{a.s.} \tag{4.7}$$

Consider for simplicity the important special case where f is the indicator function I_A of some fixed set $A \subset R$; then $f^2 = f$. We also assume $f(\cdot) = f^2(\cdot) \to c$ (C_1), that is, $|A \cap (0, n)| \sim cn$ as $n \to \infty$, where $|\cdot|$ denotes Lebesgue measure. Then (4.6) demands $c > 0$. Now

$$n^{-1/2}\left|\sum_1^{N(n)} I_A(S_k) - \sum_1^n I_A(S_k)\right| \leqslant n^{-1/2}|N_n - n|. \tag{4.8}$$

By the CLT of renewal theory [**16**, **II**, XI, §5], $n^{-1/2}(N_n - n)$ converges to normality, so the left-hand side of (4.8) is stochastically bounded. Also

$$\begin{aligned} &n^{-1/2}\left\{|A \cap (0, n)| - \sum_1^n P(S_k \in A)\right\} \\ &\qquad = n^{-1/2}E\left\{\sum_1^{N(n)} f(S_k) - \sum_1^n f(S_k)\right\} = O(1) \end{aligned} \tag{4.9}$$

as in the proof of (2.6.9) above. Combining (4.5), (4.8) and (4.9), we merely obtain the conclusion that

$$n^{-1/2}\sum_1^n \{I_A(S_k) - P(S_k \in A)\}, \qquad n = 1, 2, \ldots, \text{ is tight.} \tag{4.10}$$

Similarly, (4.5) and (4.8) alone yield that

$$n^{-1/2}\left\{\sum_1^n I_A(S_k) - |A \cap (0, n)|\right\}, \qquad n = 1, 2, \ldots, \text{ is tight.} \tag{4.11}$$

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

Whether one can obtain both convergence in law as in (4.5), and deterministic centring as in (4.10), (4.11), remains open.

Returning to (4.7)

$$\limsup_{n\to\infty} (2cn \log\log n)^{-1/2}\left\{ \sum_1^{N(n)} I_A(S_k) - |A \cap (0, n)| \right\} = 1 \quad \text{a.s.},$$

and

$$(2n \log\log n)^{-1/2}\left|\left(\sum_1^{N(n)} - \sum_1^{n} \right) I_A(S_k)\right|$$
$$\leqslant \limsup_{n\to\infty} (2n \log\log n)^{-1/2}|N_n - n| = 1 \quad \text{a.s.},$$

by the LIL of renewal theory. Hence

$$\limsup_{n} (2n \log\log n)^{-1/2}\left\{ \sum_1^{n} I_A(S_k) - |A \cap (0, n)| \right\} \leqslant 1 + c^{1/2} \quad \text{a.s.},$$

the LIL for $I_A(S_n)$. Alternatively, using (4.9),

$$\limsup_{n} (2n \log\log n)^{-1/2} \sum_1^{n} \{ I_A(S_k) - P(S_k \in A)\} \leqslant 1 + c^{1/2} \quad \text{a.s.}$$

5. Complements.

5.1. When strong mixing ('weak ϕ-mixing') [**43**, p. 212] is strengthened to ϕ-mixing, stronger convergence results often follow. It may be of interest to know that this can happen in Theorem 4 only if the tail $\bar{F}$ of F decreases exponentially and in a somewhat regular way.

THEOREM. *Under the conditions of Theorem* 4, $\{B_n\}$ *is* ϕ*-mixing if and only if there exist constants* $\varepsilon > 0$, C *such that*

$$\bar{F}(x + y) \leqslant Ce^{-\varepsilon y}\bar{F}(x), \qquad x, y > 0.$$

Then the ϕ*-mixing coefficients satisfy* $\rho_n = O(e^{-\delta n})$ *as* $n \to \infty$, *for some* $\delta > 0$.

The condition on $\bar{F}$ is equivalent to $\mathcal{F}(x) := \bar{F}(\log x)$ having *negative* upper Matuszewska index $\alpha_{\mathcal{F}}$ (Bingham and Goldie [**7**, II]). The proof (omitted) uses 'geometric ergodicity' results of Stone [**42**] in the density case, Kingman [**24**] in the lattice case.

5.2. Were $\{S_n\}$ itself strong-mixing, the use of backward recurrence times could have been avoided altogether in §2. To see that $S_n := \Sigma_1^n X_k$ is not strong mixing (for any i.i.d. random variables X_k of finite variance) consider the events $C_m := \{S_m \leqslant m\mu\}$, where $\mu := EX_1$. It is easy to see that as $m \to \infty$ with n fixed, $P(C_m \setminus C_{m+n}) \to 0$; hence

$$P(C_m \cap C_{m+n}) - P(C_m)P(C_{m+n}) \to \tfrac{1}{2} - \tfrac{1}{2}\cdot\tfrac{1}{2} = \tfrac{1}{4}.$$

So ρ_n in (2.4.1) is at least $\frac{1}{4}$, and the requirement that $\rho_n \to 0$ is violated.

5.3. While Theorem A requires only a finite first moment for X_1, all our results need a finite moment of order greater than 2. Thus it 'should' be possible to reduce

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

our stipulated moment orders by 1. However, to do so is not possible by our method of backward recurrence times, as we shall show, and on the other hand we do not think it is at present feasible to extend Meilijson's methods to the Marcinciewicz-Zygmund cases where $\frac{1}{2} < \alpha < 1$ (it no longer suffices to approximate f uniformly by simple functions, for instance). The problem of proving the Marcinciewicz-Zygmund law for $f(S_n)$ under the 'right' moment assumptions remains open.

The reason why our methods need at least second moments finite is because of the mixing properties of the backward recurrence times. As $t \to \infty$ the distribution of B_t converges to its stationary distribution, that with density $\mu^{-1}\bar{F}(\cdot)$ on $\mathbf{R}^+$. Hence

$$\begin{aligned}&|P(B_m \leqslant 1, B_{m+n} > n+1) - P(B_m \leqslant 1)P(B_{m+n} > n+1)| \\ &\quad = P(B_m \leqslant 1)P(B_{m+n} > n+1) \to \int_0^1 \mu^{-1}\bar{F}(x)\,dx \int_{n+1}^{\infty} \mu^{-1}\bar{F}(y)\,dy\end{aligned}$$

as $m \to \infty$ with n fixed. Thus the mixing coefficient ρ_n in Theorem 4 is at least $c\int_{n+1}^{\infty} \mu^{-1}\bar{F}(y)\,dy$, where $c := \int_0^1 \mu^{-1}\bar{F}(x)\,dx > 0$. If $E(X_1^l) < \infty$ but $E(X_1^{l+\varepsilon}) = \infty$ then $\rho_n = O(n^{-(l-1)})$ but $\rho_n \neq O(n^{-(l-1+\varepsilon)})$. So the loss of one moment order occurs unavoidably at this point, the rate of r-quickness in Theorem 5 being best possible in relation to the order of the ρ_n.

5.4. A multidimensional version of Proposition 3.3 is as follows. Fix $k \geqslant 1$ and let $\xi_1, \ldots, \xi_k$ be each uniformly distributed over [0, 1], independently of each other and of $\{W_t\}_{t \geqslant 0}$. Recall that $\{W_t\}$ is a Wiener process of mean μ and variance σ^2.

THEOREM. *Assume $\mu > 0$, $\sigma^2 < \infty$. Let $\mathcal{K}_T$ denote the distribution in $\mathbf{R}^k$ of $(\mu T\xi_1, \ldots, \mu T\xi_k)$ and let $\mathcal{L}_T$ denote the distribution of $(W(T\xi_1), \ldots, W(T\xi_k))$. Then for every $\varepsilon > 0$,*

$$\|\mathcal{L}_T - \mathcal{K}_T\| = o(T^{-1/2+\varepsilon}), \qquad T \to \infty.$$

(Define variation norms as in $\mathbf{R}$, by $\|\nu\| := \frac{1}{2}\int_{\mathbf{R}^k} |\nu|(d\mathbf{x})$, for a signed measure ν on the Borel sets of $\mathbf{R}^k$.) The proof is an extension of that of Proposition 3.3, using order statistics.

5.5. The methods of the paper allow a generalisation to the case of a *random* bounded function f, in place of the fixed bounded f considered so far. For instance, in the LLT-II case (S_n integer valued), suppose one has a sequence of uniformly bounded random variables $Z_1, Z_2, \ldots,$ independent of each other and of $\{S_n\}_{n \geqslant 0}$. Then if $\mathcal{P}$ satisfies the conditions of Theorem 2′ we can obtain, for r, α as in that theorem,

$$n^{-\alpha}\left\{\sum_{k=1}^{n} Z_{S_k} - \mu^{-1}\sum_{k=1}^{[\mu n]} EZ_k\right\} \to 0, \quad r\text{-quickly}.$$

To proceed generally, suppose we have a family $\{\mathcal{Q}_x\}$ of probability distributions on $\mathbf{R}$, where x ranges over the union of the supports of all the $\mathcal{P}_n$. The supports of all the $\mathcal{Q}_x$ are contained in some fixed finite interval $[-M, M]$, $M < \infty$, and we write $m(x)$ for the mean of the distribution $\mathcal{Q}_x$. When x takes only integer values

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

(the LLT-II case) we set $m(x) := m(k)$ for $k \leqslant x < k + 1$. Random variables V_1, $V_2, \ldots$ are obtained as follows: given $(S_1, S_2, \ldots) = (s_1, s_2, \ldots)$ the random variables $V_1, V_2, \ldots$ are to be (conditionally) independent, with V_k having probability distribution $\mathcal{Q}_{s_k}$, for $k = 1, 2, \ldots$. Theorems 1, 2 and 2′ remain in force, their conclusions (1.2), (1.3) and (1.4) becoming respectively

$$n^{-\alpha} \sum_{k=1}^{n} \{V_k - EV_k\} \to 0, \quad r\text{-quickly}, \tag{1.2′}$$

$$n^{-1} \sum_{k=1}^{n} EV_k - n^{-1} \int_0^n m(\mu x)\, dx = o(n^{-\delta/2+\varepsilon}), \qquad n \to \infty, \tag{1.3′}$$

$$n^{-\alpha} \left\{ \sum_{k=1}^{n} V_k - \int_0^n m(\mu x)\, dx \right\} \to 0, \quad r\text{-quickly}. \tag{1.4′}$$

To extend the proofs to this setting it suffices to augment the filtrations $\{\mathcal{F}_t\}$ and $\{\mathcal{G}_t\}$ in §2 and to take an extra expectation with respect to the relevant $\mathcal{Q}_x$ distribution in §3. Thus define

$$\mathcal{F}_t := \sigma\{N_u, V_{N_u} : 0 \leqslant u \leqslant t\}, \quad Y_k := \sum_{N_{k-1} < j \leqslant N_k} V_j$$

in §2.1, and in §3 define $V_t := V_{[t]}$, $\lambda_T := T^{-1} \int_0^T m(\mu t)\, dt$, $\mu_T := T^{-1} \int_0^T V_t\, dt$, $\nu_T := T^{-1} \int_0^T \tilde{V}_t\, dt$, where conditional on $\{W_t\}_{t \geqslant 0}$, $\tilde{V}_t$ has distribution $\mathcal{Q}_{W_t}$, only the expectation being needed.

5.6. For any i.i.d. random variables $X_1, X_2, \ldots,$ without moment assumptions, the Hewitt-Savage zero-one law shows that any almost-sure limit of $n^{-1} \Sigma_1^n f(S_k)$ must be degenerate. By the usual subsequence argument, any in-probability limit of $n^{-1} \Sigma_1^n f(S_k)$ is also degenerate. But there is no general conclusion possible about degeneracy of limit *laws*, for by Theorem A there are only degenerate limit laws possible under the conditions of that theorem, whereas the following result of Davydov and Ibragimov gives nondegenerate limit laws in a wide variety of $\mu = 0$ cases.

THEOREM B. *Assume* (i) LLT-I *or* LLT-II; (ii) $\mu = 0$ *and* $\sigma^2 < \infty$. *Then*

(a) **[14]** $Ef(S_n) \to c$ $(n \to \infty)$ *if and only if* $(2x)^{-1} \int_{-x}^{x} f(y)\, dy \to c$ $(x \to \infty)$;

(b) **[12]** *if* $x^{-1} \int_0^x f(y)\, dy \to p$, $x^{-1} \int_{-x}^0 f(y)\, dy \to q$ $(x \to \infty)$, *then* $n^{-1} \Sigma_0^n f(S_k)$ *converges in law, as* $n \to \infty$, *to* $pX + q(1 - X)$, *where* X *has the arcsine distribution on* $[0, 1]$;

(c) **[13**, §6] *if* f *vanishes on* $(-\infty, 0)$ *and* $n^{-1} \Sigma_0^n f(S_k)$ *converges in law, then* $\lim_{x \to \infty} x^{-1} \int_0^x f(y)\, dy$ *exists.*

Note that when f vanishes on $(-\infty, 0)$, Theorem B gives the equivalence of

$f(x) \to p$ (C_1), $x \to \infty$,

$Ef(S_n) \to \frac{1}{2}p$, $n \to \infty$,

$n^{-1} \Sigma_0^n f(S_k)$ converges in law,

and that the limit law is then pX, with X as above. This special case does not imply the theorem. But when $\mu > 0$, by contrast, the limit behaviour of $f(S_n)$ involves f only at $+\infty$, so it is then no essential loss to assume f vanishes on $(-\infty, 0)$. When

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

$\mu > 0$ we know by Corollary 1.1 that Cesàro convergence of f is equivalent to Cesàro convergence of $Ef(S_n)$, which is weaker than convergence of $Ef(S_n)$. What, then, is the property of f equivalent to convergence of $Ef(S_n)$? The answer is in the next result. Indicate convergence under the Borel summability method by (B) and under the Euler method of parameter p by (E_p) (see e.g. **[18]**).

THEOREM C (BINGHAM **[5]**; SEE ALSO **[6]**). *Assume* (i) LLT-I *or* LLT-II; (ii) $\mu > 0$, $\sigma^2 < \infty$, *and that f vanishes on* $(-\infty, 0)$. *Then the following are equivalent*:

$Ef(S_n) \to c \ (n \to \infty)$;

$(\varepsilon\sqrt{x})^{-1}\int_x^{x+\varepsilon\sqrt{x}} f(y)\,dy \to c \ (x \to \infty)$ *for some* (*all*) $\varepsilon > 0$;

$x^{-1}\int_0^x f(y)\,dy = c + o(x^{-1/2}) + x^{-1}\int_0^x \varepsilon(y)\,dy$, $\varepsilon(x) \to 0$ *as* $x \to \infty$;

$f(n) \to c$ (B) [*lattice case only*];

$f(n) \to c$ (E_p) *for some* (*all*) $p \in (0, 1)$ [*lattice case only*].

5.7. Many results are known on limit behaviour of functions $f(S_n)$ of random walks or Markov chains $\{S_n\}$, in a variety of contexts. Thus, in the driftless case $\mu = 0$ of Theorem A, $\{S_n\}$ is attracted (without centring) to a Brownian motion, and the limiting arcsine law is generated by this Brownian motion (cf. Billingsley **[4]**). The general theme of convergence of functions of random walks, etc., to the corresponding functions of limiting diffusions has been treated by many authors; cf. Skorohod and Slobodenyuk **[40]** (who obtain parts of Theorem A), Portenko and Prokopenko **[34]**, Kulinič **[25]**, Taraskin **[46]**. Baxter and Brosamler **[2]** consider a.s. limits of $\int_0^t f(X_s)\,ds$ for diffusions X on *compact* metric spaces.

For *integrable* f, the a.s. convergence of $n^{-1}\Sigma_0^n f(S_k)$ is dealt with by Kallianpur and Robbins **[22]**, and one can also handle it by random ergodic theorems (cf. e.g. Révész **[36]**).

When $\{S_n\}$ is a *recurrent* Markov chain one may be able to handle $\{f(S_n)\}$ by using the regenerative property at returns to a recurrent state (the 'Döblin trick': cf. Chung **[11]**, Stout **[43**, p. 325]).

Functions of certain weakly dependent random variables have been treated, under suitable mixing assumptions, etc., by Philipp and Stout **[33]**, McLeish **[29]**. Functions of strongly dependent Gaussian sequences are considered by Taqqu **[45]** and Lai and Stout **[28]**.

5.8. The strong law (in, say, the form (I) iff (IV) in Corollary 1.1) amounts formally to using the ordinary strong law $S_k/k \to \mu$ a.s. to pass from $n^{-1}\Sigma_1^n f(S_k) \to c$ a.s. to $n^{-1}\int_0^n f(\mu y)\,dy \to c$. Analytically, this type of operation resembles the second consistency theorem for Riesz (typical) means (see e.g. Chandrasekharan and Minakshisundaram **[8]**). From a numerical analysis viewpoint, this is a procedure of Monte Carlo type. Suppose, for instance, we know a bounded function f of continuous argument is Cesàro-summable to c, and we wish to calculate c numerically. An alternative to using the approximation $x^{-1}\int_0^x f \sim c$ for large x and numerical integration would be to simulate a random walk S_n and use $n^{-1}\Sigma_1^n f(S_k) \sim c$ for large n (cf. e.g. Niederreiter **[31]**). More generally, for any bounded measurable f, Theorem 2′ shows that the value of $\int_0^n f(\mu t)\,dt$ can be approximated with error $o(n^\alpha)$ by summing the values of f at the n simulated points $S_1, \ldots, S_n$.

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

In the same vein, $\Sigma_0^n f(k + U_k)$ (U_k independent and uniform on [0, 1]) is easily seen to have as good convergence properties as $\Sigma_1^n f(S_k)$ (cf. e.g. Petrov [**32**, Theorem 12 p. 272]; r-quickness results follow from our Theorem 5).

NOTE ADDED IN PROOF. Perhaps the best interpretation, told us by I. Meilijson (unpublished), is in terms of *random quality control*. The probability p of an article being defective is to be estimated. Taking $f := I_D$, with $D \subset \mathbf{N}$ the set of defectives, one can sample at the instants of a random walk $\{S_n\}$, and then $n^{-1}\Sigma_1^n I_D(S_k)$ is a strongly consistent estimator for p, with error $o(n^{1-\alpha})$.

REFERENCES

1. L. B. Baum and M. Katz, *Convergence rates in the law of large numbers*, Trans. Amer. Math. Soc. **120** (1965), 108–123.

2. J. R. Baxter and G. R. Brosamler, *Energy and the law of the iterated logarithm*, Math. Scand. **38** (1976), 115–136.

3. H. C. P. Berbee, *The equilibrium random walk*, Report 45, Wiskundig Seminarium der Vrije Universiteit, Amsterdam, 1976.

4. P. Billingsley, *Convergence of probability measures*, Wiley, New York, 1968.

5. N. H. Bingham, *Tauberian theorems and the central limit theorem*, Ann. Probab. **9** (1981), 221–231.

6. _____, *Tauberian theorems for Borel and related summability methods* (to appear).

7. N. H. Bingham and C. M. Goldie, *Extensions of regular variation*, I. *Uniformity and quantifiers*, II. *Representations and indices*, Proc. London Math. Soc. (to appear).

8. K. Chandrasekharan and S. Minakshisundaram, *Typical means*, Oxford Univ. Press, London, 1952.

9. Y. S. Chow and T. L. Lai, *Some one-sided theorems on the tail distribution of sample sums with applications to the last time and largest excess of boundary crossings*, Trans. Amer. Math. Soc. **208** (1975), 51–72.

10. Y. S. Chow and H. Teicher, *Probability theory*, Springer-Verlag, Berlin, 1978.

11. K. L. Chung, *Markov chains with stationary transition probabilities*, 2nd ed., Springer-Verlag, Berlin, 1976.

12. Yu. A. Davydov, *Limit theorems for functionals of processes with independent increments*, Theory Probab. Appl. **18** (1973), 431–441.

13. _____, *Sur une classe de fonctionelles des processus stables et des marches aléatoires*, Ann. Inst. H. Poincaré (B) **10** (1974), 1–29.

14. Yu. A. Davydov and I. A. Ibragimov, *On asymptotic behaviour of some functionals of processes with independent increments*, Theory Probab. Appl. **16** (1971), 162–167.

15. J. L. Doob, *Stochastic processes*, Wiley, New York, 1953.

16. W. Feller, *An introduction to probability theory and its applications*. vol. I, 3rd ed.; vol. II, 2nd ed., Wiley, New York, 1968; 1971.

17. A. Gut, *On the moments and limit distributions of some first-passage times*, Ann. Probab. **2** (1974), 277–308.

18. G. H. Hardy, *Divergent series*, Oxford Univ. Press, London, 1949.

19. I. A. Ibragimov, *Some limit theorems for stationary processes*, Theory Probab. Appl. **7** (1962) 349–382.

20. I. A. Ibragimov and Yu. V. Linnik, *Independent and stationary sequences of random variables*, Wolters-Noordhoff, Groningen, 1971.

21. N. L. Johnson and S. Kotz, *Distributions in statistics: continuous univariate distributions*. I, Wiley-Interscience, New York, 1970.

22. G. Kallianpur and H. Robbins, *The sequence of sums of independent random variables*, Duke Math. J. **21** (1954), 285–307.

23. S. Karlin, *On the renewal equation*, Pacific J. Math. **5** (1955), 229–257.

24. J. F. C. Kingman, *Regenerative phenomena*, Wiley, New York, 1972.

25. G. L. Kulinič, *On the asymptotic behaviour of the distribution of functionals of the type* $\int_0^t g(\xi(s))\,ds$ *of diffusion processes*, Teor. Verojatnost. i Mat. Statist. **8** (1973), 99–105. (Russian).

26. T. L. Lai, *On r-quick convergence and a conjecture of Strassen*, Ann. Probab. **4** (1976), 612–627.

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

27. _____, *Convergence rates and r-quick versions of the strong law for stationary mixing sequences*, Ann. Probab. **5** (1977), 693–706.

28. T. L. Lai and W. F. Stout, *Limit theorems for sums of dependent random variables*, Z. Wahrsch. Verw. Gebiete **51** (1980), 1–14.

29. D. L. McLeish, *A maximal inequality and dependent strong laws*, Ann. Probab. **3** (1975), 829–839.

30. I. Meilijson, *The average of the values of a function at random points*, Israel J. Math. **15** (1973), 193–203.

31. H. Niederreiter, *Quasi-Monte Carlo methods and pseudorandom numbers*, Bull. Amer. Math. Soc. **84** (1978), 957–1041.

32. V. V. Petrov, *Sums of independent random variables*, Springer-Verlag, Berlin, 1975.

33. W. Philipp and W. F. Stout, *Almost sure invariance principles for partial sums of weakly dependent random variables*, Mem. Amer. Math. Soc. No. 161 (1975).

34. N. I. Portenko and L. N. Prokopenko, *A method of proving limit theorems for functionals of random walks*, Theory of Random Processes **1** (1973), 107–119. (Russian)

35. N. U. Prabhu, *Stochastic processes*, Macmillan, New York, 1965.

36. P. Révész, *Some remarks on the random ergodic theorem*. I, II, Publ. Inst. Math. Hungar. Acad. Sci. **5** (1960), 375–381; **6** (1961), 205–213.

37. M. Rosenblatt, *A central limit theorem and a strong mixing condition*, Proc. Nat. Acad. Sci. U.S.A. **42** (1956), 43–47.

38. R. J. Serfling, *Moment inequalities for the maximum cumulative sum*, Ann. Math. Statist. **41** (1970), 1227–1234.

39. S. H. Sirazhdinov and M. Mamatov, *On convergence in the mean for densities*, Theory Probab. Appl. **7** (1962), 433–437.

40. A. V. Skorohod and N. P. Slobodenyuk, *Limit theorems for random walks*, Izdat Naukova Dumka, Kiev, 1970. (Russian)

41. C. Stone, *On characteristic functions and renewal theory*, Trans. Amer. Math. Soc. **120** (1965), 327–342.

42. _____, *On absolutely continuous components and renewal theory*, Ann. Math. Statist. **37** (1966), 271–275.

43. W. F. Stout, *Almost sure convergence*, Academic Press, New York, 1974.

44. V. Strassen, *Almost sure behaviour of sums of independent random variables and martingales*, Proc. Fifth Berkeley Sympos. Math. Statist. Prob., vol. 2, part 1, Univ. of California Press, Berkeley, 1967, pp. 315–343.

45. M. S. Taqqu, *Law of the iterated logarithm for sums of nonlinear functions of Gaussian variables that exhibit a long-range dependence*, Z. Wahrsch. Verw. Gebiete **40** (1977), 203–238.

46. A. F. Taraskin, *Certain limit theorems for stochastic integrals*, Theory of Random Processes **1** (1973), 119–133. (Russian)

Department of Mathematics, Westfield College, University of London, London NW3 7ST, United Kingdom

School of Mathematical and Physical Sciences, University of Sussex, Falmer, Brighton BN1 9QH, United Kingdom

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use