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 \( \mu \) 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 \( n^{-\alpha}(\sum_1^n f(S_k) - \int f(\mu) \, d\mu) \to 0 \) (\( r \)-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 \( F, \mu, \sigma^2 \) for the distribution function, mean and variance of \( X_1 \). Setting \( S_n = X_1 + \cdots + X_n, S_0 = 0 \), we write \( P, \mathbb{P}, \mathbb{P}_n \) 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 \( \mathbb{P}_k \) has an absolutely continuous component.

LLT-I' \( \mathbb{P} \) has an absolutely continuous component.

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

Our results in the LLT-II case may be re-expressed for any lattice law \( \mathbb{P} \), 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 \([k, k+1), k \in \mathbb{Z} \). 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

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

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.
(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 \( \sup\{ n > 1: |U_n| > \varepsilon \} \) has finite \( r \)th moment (by convention, \( \sup \emptyset = 0 \)). 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) \( \mu > 0, E\{X_1^+\} < \infty \) for some \( l > 2 \), \( E\{X_1^-\} < \infty \) for some \( \lambda > 2 \).

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

\[
0 < r < \min\{\left(\alpha - \frac{1}{2}\right)\lambda, \alpha(l - 1) - 1\}.
\]

Then

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

Note the trade-off between the values of \( \alpha \) and \( r \); the smaller \( \alpha \), the stronger the statement that \( n^{-\alpha} \sum \{ f(S_k) - Ef(S_k) \} \to 0 \), 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 \( \delta \) satisfying \( 0 < \delta < 1 \).

Then for each \( \varepsilon > 0 \),

\[
n^{-1} \sum_{k=1}^{n} Ef(S_k) - n^{-1} \int_{0}^{\varepsilon} f(\mu x) \, dx = o(n^{-\delta/2+\varepsilon}), \quad n \to \infty.
\]

Combining Theorems 1 and 2 we have

**Theorem 2' (r-quick equiconvergence).** Assume

(i) LLT-I' or LLT-II;
(ii) \( \mu > 0, \) and \( E\{X_1^+\} < \infty, E\{X_1^-\} < \infty \) for some \( l > 2, \lambda > 2 \).

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

\[
n^{-\alpha} \left\{ \sum_{k=1}^{n} f(S_k) - \int_{0}^{\varepsilon} f(\mu x) \, dx \right\} \to 0, \quad \text{r-quickly}.
\]
In the important special case when $E(|X|^3) < \infty$ 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,) 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) $n^{-1} \sum_{i=1}^{n} Ef(S_k) \to c$;

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

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

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

As $f$ is bounded, I is equivalent to

$$n^{-1} \int_{0}^{n} f(\mu x) \, dx = (n\mu)^{-1} \int_{0}^{n\mu} f(y) \, dy \to c \quad (n \in \mathbb{Z}, \to \infty),$$

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, $E(|X|^{2+\delta}) < \infty$ 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 $\mu$ 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 ($\{S_n\} \text{ Poisson}$) 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.
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 $\mathbb{P}$ 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 $N_t := \max\{n: S_n < t\}$ is a renewal process. Write $Y_k := \sum_{N_{k-1} < i < N_k} f(S_i)$. Then the results of §2.3 on $r$-quick convergence for renewal processes are used to show (1.1) equivalent to

$$n^{-a} \sum_{k=1}^{n} (Y_k - EY_k) \rightarrow 0, \ r\text{-quickly.}$$

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

$$n^{-a} \sum_{k=1}^{n} (E(Y_k | \mathbb{F}_{k-1}) - EY_k) \rightarrow 0, \ r\text{-quickly.}$$

Let $B_t := t - S(N_t)$ be the spent lifetime (backward recurrence time) at time $t > 0$. We show that the above can be rewritten as $n^{-a} \sum_{k=1}^{n} (g_k(B_k - 1) - Eg_k(B_k - 1)) \rightarrow 0$ $r$-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 $U_n \rightarrow 0$, $r$-quickly, and $U'_n \rightarrow 0$, $r'$-quickly, then $U_n + U'_n \rightarrow 0$, $(r \land r')$-quickly.

We leave the proof as an exercise.

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

$$\int_0^\infty t^{r-1} P\left( \sup_{k \geq t} |U_k| > \epsilon \right) dt < \infty \quad \text{for all} \ \epsilon > 0.$$

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

LEMMA 2.2.3. Assume the martingale difference sequence $\{\xi_n\}$ satisfies $E(|\xi_n|^r) < M$, $n = 1, 2, \ldots$, where $M$ and $r > 2$ are constants. Then for every $\alpha, r$ satisfying $\alpha > \frac{1}{2}$, $0 < r < (\alpha - \frac{1}{2})v$, $n^{-\alpha} \sum_{k=1}^{n} \xi_k \rightarrow 0$, $r$-quickly.

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

Proof. For some constant $K = K_v$,

$$E\left(\sum_{i=1}^{n} |\xi_i|^{v}\right) < Kn^{v/2}, \quad n = 1, 2, \ldots,$$

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

$$P\left(\max_{k=1,\ldots,n} \left| \sum_{i=1}^{k} \xi_i \right| > en^a \right) = P\left(\max_{k=1,\ldots,n} \left| \sum_{i=1}^{n} \xi_i \right| > (en^a)^{v}\right)$$

by Doob’s inequality [15, p. 314], since $|\sum_{i=1}^{n} \xi_i|^v$ is a nonnegative $L_1$ submartingale. Thus

$$\sum_{n=1}^{\infty} n^{-1} P\left(\max_{k=1,\ldots,n} \left| \sum_{i=1}^{k} \xi_i \right| > en^a \right) < \sum_{n=1}^{\infty} n^{-1} (en^a)^{-r} Kn^{v/2} < \infty.$$

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 $\alpha < 1$. In the former case much more than $r$-quick convergence is true, without conditions:

Theorem 3a. For each $\alpha > 1$, $\epsilon > 0$ the random variable

$$T_{\alpha,\epsilon} = \sup\{t > 0: t^{-\alpha}N_t > \epsilon\}$$

has tail behaviour

$$P(T_{\alpha,\epsilon} > t) < ae^{-bt^\alpha}, \quad all \ t > 0,$$

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

Proof. Let $[\cdot]$ denote integer part. On the event $\{T_{\alpha,\epsilon} > t\}$ we have $N_u > eu^\alpha$ for some $u > t$; hence $S([eu^\alpha] + 1) < u$ and so $S_k < (k/\epsilon)^{1/\alpha}$ where $k = [eu^\alpha] + 1 > et^\alpha$. 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

$$2nm - S'_n = \sum_{i=1}^{n} \min(X_i, 1) < S_n, \quad n = 1, 2, \ldots.$$

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

$$P(T_{\alpha,\epsilon} > t) < P(S'_k > mk) \leq \sum_{k > et^\alpha} P(S'_k > mk).$$

Because the $X'_k$ are i.i.d. with $|X'_k| < 1$, $EX'_k = 0$, we may use Prohorov’s exponential bound. Set $\nu = \var X'_1$. By [43, Theorem 5.2.2(ii)],

$$P(S'_n > mn) < \exp\left(\frac{-1}{2} mn\ \text{arcsinh}\left(\frac{1}{2} m/\nu\right)\right) = e^{-Kn},$$

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),

$$P(T_{\alpha, e} \geq t) \leq (1 - e^{-K})^{-1}\exp\left\{-K\left(\lceil \alpha t \rceil + 1\right)\right\}$$

$$\leq (1 - e^{-K})^{-1}\exp(-K\alpha t),$$

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

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

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

if and only if

(2.3.4) $E(X_{1}^{l}) < \infty$ and $EX_{1} = m$.

**Proof.** By Lemma 2.2.2, (2.3.3) is equivalent to

(2.3.5) $\int_{0}^{\infty} u^{-\alpha/2}P\left(\sup_{t \geq u} t^{-\alpha}|N_{t} - t/m| > \varepsilon\right) du < \infty$, all $\varepsilon > 0$.

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

(2.3.6) $\int_{0}^{\infty} u^{-\alpha/2}P\left(\sup_{k \geq u} k^{-\alpha}|S_{k} - km| > \varepsilon\right) du < \infty$, all $\varepsilon > 0$.

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: $S_{n} = \sup\{t: N_{t} < n\}$, $N_{t} = \max\{n: S_{n} < t\}$, 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

(2.37) $\int_{0}^{\infty} u^{-\alpha/2}P\left(\sup_{t \geq u} t^{-\alpha}|N_{t} - t| > \varepsilon\right) du$

$$\leq \int_{0}^{\infty} u^{-\alpha/2}P\left(\sup_{t \geq u} t^{-1/2}N_{t} - 1\right) > \frac{1}{2}\right) du$$

$$+ \int_{0}^{\infty} u^{-\alpha/2}P\left(\sup_{t \geq u} t^{-\alpha}|N_{t} - t| \geq \varepsilon, \sup_{t \geq u} t^{-1/2}N_{t} - 1 \leq \frac{1}{2}\right) du.$$
by (2.3.6) (with \( m = 1 \)). As for the other term on the right of (2.3.7), if \( v^{-a}|N_v-v| > \varepsilon \) for some \( v > u \) and \( |r_1^{-1}N_t-1| < \frac{1}{2} \) for all \( t > u \) then \( |N_v-v| > \varepsilon (\frac{2}{3})^a N_v^a \) and \( N_u > \frac{1}{2} u \). Setting \( k = N_v \) we find \( S_k < v < S_{k+1} \) so that one of the inequalities

\[
|k - S_k| > \varepsilon \left( \frac{2}{3} \right)^a k^a, \quad |k - S_{k+1}| > \varepsilon \left( \frac{2}{3} \right)^a k^a
\]

must hold. Also \( k > N_u > \frac{1}{2} u \). Thus

\[
\int_0^\infty u^{a-2} \left( \sup_{t > u} |t^{-a}|N_t-t| > \varepsilon, \sup_{t > u} |t^{-1}N_t-1| < \frac{1}{2} \right) du
\]

\[
< \int_0^\infty u^{a-2} \left( \sup_{k > u/2} k^{-a}|k - S_k| > \varepsilon \left( \frac{2}{3} \right)^a \right) du
\]

\[
+ \int_0^\infty u^{a-2} \left( \sup_{k > u/2} k^{-a}|k - S_{k+1}| > \varepsilon \left( \frac{2}{3} \right)^a \right) du
\]

\[
= 2^{a-1} \int_0^\infty u^{a-2} \left( \sup_{k > u} k^{-1}|k - S_k| > \varepsilon \left( \frac{2}{3} \right)^a \right) du
\]

\[
+ 2^{a-1} \int_0^\infty u^{a-2} \left( \sup_{k > u} k^{-a}\varepsilon (\frac{2}{3})^a \right) du.
\]

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 \( F_t := S(N_t + 1) - t \) for the forward recurrence time or remaining lifetime.

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

\[
|P(C \cap D) - P(C)P(D)| \leq \rho_n.
\]

THEOREM 4. Assume LLT-I or LLT-II, and suppose \( E(X_1^l) < \infty \) for some \( l > 2 \). 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),

\[
|P(D | F_m = x) - P(D)| \leq cn^{-(l-1)}.
\]

PROOF. First assume LLT-I. Write \( \mathcal{B}, \mathcal{B}^\infty \) for the Borel \( \sigma \)-fields in \( \mathbb{R}, \mathbb{R}^\infty \). Then

\[
D = \{ \omega: (B_{m+n}(\omega), B_{m+n+1}(\omega), \ldots) \in D^0 \}
\]

for some \( D^0 \in \mathcal{B}^\infty \). The transition probability

\[
Q(y, E) := P((B_{m+n+1}, B_{m+n+2}, \ldots) \in E | B_{m+n} = y),
\]

\( y > 0, E \in \mathcal{B}^\infty \),
is well defined. Write $U = \sum_{k=0}^{\infty} F^{(k)}$ 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])

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

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

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

We define

\[(2.4.7) \quad P_x(D^0) := \int \int_{\{(y, z) \in D^0\}} Q(y, dz) \bar{F}(y) \mu^{-1} dy\]

(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 $\mu^{[l]}_y$ 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 $r(x) = \int_0^\infty (y - x)F(dy)$ for $x > 0$. Since $E(X^l) < \infty$,

\[r(x) < \int_0^\infty yF(dy) < x^{-[l]} \int_0^\infty y^lF(dy) = o(x^{-(l-1)}).\]

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

\[(2.4.8) \quad p(x) = 1/\mu + q(x), \quad x \in \mathbb{R},\]

where

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

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

\[(2.4.10) \quad P(D | F_m = x) - P_x(D^0) = \int \int_{\{(y, z) \in D^0\}} Q(y, dz) \bar{F}(y) \{ U_2(n - x - dy) + q(n - x - y) dy \}.\]

Write $\mu_{\lfloor \ast \rceil} = E(X^\lfloor \ast \rceil)$:

\[(2.4.11) \quad 0 \leq \int \int_{\{(y, z) \in D^0\}} Q(y, dz) \bar{F}(y) U_2(n - x - dy) \leq \int_0^\infty \int_{\mathbb{R}^\infty} \bar{F}(y) U_2(n - x - dy)\]

\[\leq \int_0^{n/4} U_2(n - x - dy) + \bar{F}(n/4) \int_{n/4}^{n/4} U_2(n - x - dy)\]

\[\leq U_2[n/4, \infty) + \bar{F}(n/4) U_2[0, \infty)\]

\[\leq (4/n)^{[l]} \mu^{[l]} \mu + (4/n)^{[l]} \mu U_2[0, \infty).\]
On the other hand, writing \( q_0 := \sup_{t > 0} |t^{-1} q(t)|, \) \( q_1 := \sup_{t \in \mathbb{R}} |q(t)|, \) both finite by (2.4.9) and continuity,

\[
\left| \int \int \{ (y, z) \in D^0 \} Q(y, dz) \bar{F}(y) q(n - x - y) \, dy \right|
\leq \int_0^\infty \bar{F}(y) |q(n - x - y)| \, dy
\leq \int_0^{n/4} \bar{F}(y)(4/n)^{-1}(n - x - y)^{-1} |q(n - x - y)| \, dy
\]

(2.4.12)

\[
+ \int_{n/4}^\infty \bar{F}(y) |q(n - x - y)| \, dy
\leq (4/n)^{-1} q_0 \int_0^{n/4} \bar{F}(y) \, dy + q_1 \int_{n/4}^\infty \bar{F}(y) \, dy
\leq (4/n)^{-1} q_0 \mu + q_1 \int_{n/4}^\infty (x - n/4) F(dx)
\leq (4/n)^{-1} q_0 \mu + (4/n)^{-1} \mu q_1.
\]

From (2.4.10)–(2.4.12), we thus have

(2.4.13) \[|P(D | F_m = x) - P_s(D^0)| \leq \frac{1}{2} c' n^{-(t-1)}\]

for some constant \( c' \). Now

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

so from (2.4.4), \[P(D) = \int \int \{ (y, z) \in D^0 \} Q(y, dz) \bar{F}(y) U(m + n - dy), \]

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

\[|P(D) - P_s(D^0)| \leq \frac{1}{2} c' n^{-(t-1)}\]

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 \( u_n = 1/\mu + o(n^{-(t-1)}) \) (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 \( C \cap \{ F_m \leq \frac{1}{2} n \} \), 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 \( R(db) \), \( b = (b_0, \ldots, b_m) \). Given \( B_m = b_m > 0 \), the conditional distribution of \( F_m \) is

\[P(F_m \in A | B_m = b_m) = \int_A F(b_m + dx) / \bar{F}(b_m), \quad A \in \mathfrak{B}, b_m > 0.\]

So

\[P(C_n) P(D) = P(D) \int_{C^0} \int_0^{n/2} \left\{ F(b_m + dx) / \bar{F}(b_m) \right\} R(db),\]

\[P(C_n \cap D) = \int_{C^0} \int_0^{n/2} P(D | F_m = x) \left\{ F(b_m + dx) / \bar{F}(b_m) \right\} R(db).\]
Subtract and use the lemma:
\[ |P(C_n \cap D) - P(C_n)P(D)| \leq c_n^{-1} \int_{c=0}^{n/2} \left\{ \int_{0}^{\infty} \frac{F(b_m + dx)}{\bar{F}(b_m)} R(d\theta) \right\} R(d\theta) \]
\[ \leq c_n^{-((l-1))} P(C_n) \leq c_n^{-(l-1)}. \]

To remove the truncation of \( F_m \), use
\[ P(C \cap D) = P(C_n \cap D) + P(C \cap \{ F_m > \frac{1}{2}n \} \cap D), \]
\[ P(C) = P(C_n) + P(C \cap \{ F_m > \frac{1}{2}n \}), \]

obtaining
\[ |P(C \cap D) - P(C)P(D)| \leq |P(C_n \cap D) - P(C_n)P(D)| + 2P(F_m > \frac{1}{2}n) \]
\[ \leq c_n^{-(l-1)} + 2P(F_m > \frac{1}{2}n). \]

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 \( p_0 := \sup_{i>0} P(i) < \infty \))
\[ P(F_m > \frac{1}{2}n) = \int_{0}^{m} \bar{F}(\frac{1}{2}n + m - x) U(\pi) U(dx) + \int_{0}^{m} \bar{F}(\frac{1}{2}n + m - x) p(x) dx \]
\[ \leq \mu_l(2/n)U(\pi) U[0, \infty) + p_0 \int_{n/2}^{\infty} \bar{F}(x) dx \]
\[ \leq \mu_l(2/n)U(\pi) U[0, \infty) + p_0(2/n)^{l-1} \mu_l \leq c_1 n^{-(l-1)}, \]

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 \( (\eta_1 + \cdots + \eta_n)/n^a \), where the \( \eta_k \) 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 \( g_k: \mathbb{R} \to \mathbb{R} \) 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 \( |\eta_k| \leq M \) for all \( k \), where \( M \) is constant. Let \( \alpha, p \) be any real numbers satisfying \( \alpha > \frac{1}{2}, \theta > p > 1/\alpha \). If \( \alpha < 1 \) assume \( E\eta_k = 0 \) for all \( k \). Then
\[ (\eta_1 + \cdots + \eta_n)/n^a \to 0, \quad (p\alpha - 1)-quickly. \]

**Proof.** Set \( V_0 = 0, V_n = \eta_1 + \cdots + \eta_n, n \geq 1 \). When \( \alpha > 1 \) we have \( |EV_n/n^a| \leq Mn^{1-a} \to 0 \); 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
assume $E\eta_k = 0$ for all $k$. By [27, Lemma 4] it suffices to prove

$$\sum_{n=1}^{\infty} n^{p\alpha-2} P\left( \max_{j=1,\ldots,n} |V_j| > \varepsilon n^\alpha \right) < \infty \quad \text{for all } \varepsilon > 0.$$  \hspace{1cm} (2.5.1)

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

$$\sum_{n=1}^{\infty} n^{p\alpha-1} \rho_n' < \infty,$$  \hspace{1cm} (2.5.2)

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

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

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

$$c_n(\varepsilon) := \sup_{r=0,1,\ldots} P\left( \max_{j=1,\ldots,n} |V_{r,j}| > \varepsilon n^\alpha \right) < 1.$$  \hspace{1cm} (2.5.4)

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

$$\max_{j=1,\ldots,n'} |V_{r+i-1,j}| < \frac{1}{2} \varepsilon n^\alpha/k.$$  \hspace{1cm} (2.5.5)

If $\tau_r = i$ then

$$\max_{j=1,\ldots,i-1} |V_{r,j}| < \frac{1}{2} \varepsilon n^\alpha/k.$$  \hspace{1cm} (2.5.6)

If also it is the case that

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

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

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

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

$$\left\{ \tau_r = i, \max_{j=1,\ldots,n} |V_{r,j}| > \varepsilon n^\alpha(1 - \nu/k) \right\}$$

$$\subset \left\{ \tau_r = i, \max_{j=1,\ldots,n} |V_{r+i-1+n',j}| > \varepsilon n^\alpha(1 - (\nu + 1)/k) \right\}.$$  \hspace{1cm} (2.5.8)

Therefore

$$P\left( \max_{j=1,\ldots,n} |V_{r,j}| > \varepsilon n^\alpha(1 - \nu/k) \right)$$

$$< \sum_{i=1}^{n} P(\tau_r = i) \max_{j=1,\ldots,n} |V_{r+i-1+n',j}| > \varepsilon n^\alpha(1 - (\nu + 1)/k)).$$  \hspace{1cm} (2.5.9)

By strong mixing, the latter sum is at most

$$\sum_{i=1}^{n} P(\tau_r = i) \left[ P\left( \max_{j=1,\ldots,n} |V_{r+i-1+n',j}| > \varepsilon n^\alpha(1 - (\nu + 1)/k) \right) + n\rho_n' \right]$$

$$< P(\tau_r < n)c_n(\varepsilon(1 - (\nu + 1)/k)) + n\rho_n'$$

$$< c_n\left( \frac{1}{2} \varepsilon/k \right) \cdot c_n(\varepsilon(1 - (\nu + 1)/k)) + n\rho_n'.$$
Substituting into (2.5.8) and taking the supremum over \( r \) gives
\[
c_n(\varepsilon(1 - \nu/k)) \leq c_n\left(\frac{1}{2} \varepsilon/k\right) \cdot c_n(\varepsilon(1 - (\nu + 1)/k)) + n\rho_n.
\]
This holds for \( \nu = 0, 1, \ldots, k - 2 \), whence
\[
c_n(\varepsilon) \leq \left\{ c_n\left(\frac{1}{2} \varepsilon/k\right)\right\}^{k-1} \cdot c_n(\varepsilon/k) + n(k - 1)\rho_n.
\]
Clearly \( c_n(\varepsilon) \) is nonincreasing in \( \varepsilon \), so we conclude
\[(2.5.9)\quad c_n(\varepsilon) \leq \left\{ c_n\left(\frac{1}{2} \varepsilon/k\right)\right\}^k + nk\rho_n.
\]
Now by Ibragimov [19, Lemma 1.2], \( |E(\eta_i \eta_i+1)| \leq 4M^2 \rho_j, \) \( i = 1, 2, \ldots, j = 0, 1, \ldots \) (where \( \rho_0 = 1 \)), whence by Serfling [38, Corollary A2],
\[
E\left\{\left(\max_{j=1,\ldots,n} |V_{ij}|\right)^2\right\} < 8M^2n(\log_2(2n))^2 \sum_0^\infty \rho_i.
\]
Chebychev’s inequality gives
\[
P\left(\max_{j=1,\ldots,n} |V_{ij}| > \frac{1}{2} \varepsilon n/k\right) \leq Kn^{-2a}(\log_2(2n))^2
\]
for \( K = 32(kM/\varepsilon)^2 \sum_i \rho_i \), and so \( c_n(\frac{1}{2} \varepsilon/k) < Kn^{-2a}(\log_2(2n))^2 \). From (2.5.9),
\[
c_n(\varepsilon) \leq K^k n^{-(2a-1)k}(\log_2(2n))^{2k} + nk\rho_n,
\]
and so \( \sum_0^\infty n^{\rho_n - 2}c_n(\varepsilon) < \infty \) 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, \( \lambda \) 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 \( N_i, B_i, F_i \) as in §§2.1, 2.4, and again write \( \mathcal{G}_i := \sigma(N_u; 0 < u < t) \). (Actually we need \( \mathcal{G}_n \) only for integer \( n \). As \( N_0 = 0, \mathcal{G}_0 = \{\varnothing, \Omega\} \). Defining \( Y_n \) as in §2.1, we see that \( \{Y_n, \mathcal{G}_n\} \) is an adapted sequence, so that \( \{Y_n - E(Y_n|\mathcal{G}_n), \mathcal{G}_n\} \) is a martingale difference sequence. Further, the conditional distribution of \( F_{n-1} \) given \( \mathcal{G}_{n-1} \) involves only \( B_{n-1} \):
\[
P(F_{n-1} > x|\mathcal{G}_{n-1}) = \bar{F}(B_{n-1} + x)/\bar{F}(B_{n-1}) \quad \text{a.s., } x > 0,
\]
(recall \( \bar{F} := 1 - F \)), from which it follows that \( E(Y_n|\mathcal{G}_{n-1}) = g_n(B_{n-1}) \) a.s., \( n = 1, 2, \ldots \), for certain measurable functions, \( g_1, g_2, \ldots \). The first lemma concerns these functions. Recall \( U := \sum_0^n F(n), \sup|f(\cdot)| \leq M \).

**Lemma 2.6.1.** \( |g_n(b)| < MU(1) < \infty \) for all \( b > 0, n = 1, 2, \ldots \).

**Proof.**
\[
|g_n(b)| = \left| E\left( \sum_{N_{n-1} < k < N_n} f(S_k)|B_{n-1} = b\right) \right|
\leq ME(N_n - N_{n-1}|B_{n-1} = b)
= ME(E(N_n - N_{n-1}|F_{n-1})|B_{n-1} = b).
\]
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 \( r, \alpha \) satisfying \( \alpha > \frac{1}{2}, 0 < r < \alpha(l - 1) - 1 \). Then

\[
(2.6.1) \quad n^{-\alpha} \sum_{1}^{n} (Y_k - EY_k) \to 0, \text{ r-quickly.}
\]

**Proof.** Pick \( \nu > \max(2, r/(\alpha - \frac{1}{2})) \). As in the previous proof, \( |Y_n| < MN'_1 \); hence \( E(|Y_n|^\nu) < M'E((N_1 + 1)^\nu) \). 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,

\[
(2.6.2) \quad n^{-\alpha} \sum_{1}^{n} \{ Y_k - g_k(B_{k-1}) \} \to 0, \text{ r-quickly.}
\]

On the other hand, consider the sequence \( g_k(B_{k-1}) - E g_k(B_{k-1}), k = 1, 2, \ldots \), 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 \( g_k(B_{k-1}) - E g_k(B_{k-1}) \). By Theorem 5, since \( r = pa - 1 \) where \( l - 1 > p > 1/\alpha \),

\[
(2.6.3) \quad n^{-\alpha} \sum_{1}^{n} \{ g_k(B_{k-1}) - E g_k(B_{k-1}) \} \to 0, \text{ r-quickly.}
\]

However \( E g_k(B_{k-1}) = E(E(Y_k|S_{k-1})) = EY_k \), 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

\[
(2.6.4) \quad n^{-\alpha} \left\{ \sum_{1}^{N(n)} f(S_k) - E \sum_{1}^{N(n)} f(S_k) \right\} \to 0, \text{ r-quickly,}
\]

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

\[
(2.6.5) \quad n^{-\alpha} \left\{ \sum_{1}^{N(n)} f(S_k) - \sum_{1}^{[n/\mu]} f(S_k) \right\} \to 0, \text{ r-quickly,}
\]

and

\[
(2.6.6) \quad n^{-\alpha} E \left\{ \sum_{1}^{N(n)} f(S_k) - \sum_{1}^{[n/\mu]} f(S_k) \right\} \to 0.
\]

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

\[
n^{-\alpha} \left| \sum_{1}^{N(n)} f(S_k) - \sum_{1}^{[n/\mu]} f(S_k) \right| \leq M n^{-\alpha} (|N_n - n/\mu| + 1)
\]
and the latter tends to zero (la — 1)-quickly for \( \frac{1}{2} < \alpha < 1 \), \( \infty \)-quickly for \( \alpha > 1 \), so 
\( r \)-quickly in either case since \( r < \alpha a - 1 \).

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

\[
\left| n^{-\alpha}E \left\{ \sum_1^{N(n)} f(S_k) - \sum_1^{[n/\mu]} f(S_k) \right\} \right| \leq Mn^{-\alpha}(E|N_n - n/\mu| + 1)
\leq Mn^{-\alpha}(E|N_n - EN_n| + |EN_n - n/\mu| + 1)
\leq Mn^{-\alpha}\{\text{var} N_n\}^{1/2} + O(1)\}
= Mn^{-\alpha}\{(\sigma^2 n/\mu^3)^{1/2}(1 + o(1)) + O(1)\} \to 0
\]
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 \( A := \sum_{k=1}^\infty n^{-1}P(S_n < 0) \). 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 \((\mathbb{R}, \mathfrak{B})\) induced by \( F \) and its absolutely continuous component by \( \nu, \nu_c \) respectively. If the support of \( \nu_c \) intersects \((0, \infty)\) then immediately

\[
\forall B \in \mathfrak{B}^+, \quad P(S(L_1) \in B) > P(X_1 \in B) > \nu_c(B),
\]
so \( G \) has a nontrivial absolutely continuous component. We therefore assume from now on that the support of \( \nu_c \) is contained in \((-\infty, 0]\). Absolute continuity means that

\[
(2.7.1) \quad \nu_c(B) = \int_B w(x) \, dx, \quad \text{all } B \in \mathfrak{B},
\]
for some nonnegative measurable function \( w \). By assumption \( \nu_c \) 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 \( 0 < a < b \), 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 \( \nu_{cr} \) denote the restriction of \( \nu_c \) to the interval \((-x_1 - \varepsilon, -x_1 + \varepsilon)\) and let \( \nu_r \) denote the restriction of \( \nu \) to the interval \((x_0 - \varepsilon, x_0 + \varepsilon)\). Both \( \nu_{cr} \) and \( \nu_r \) 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 \( \{S_k\}_{k=0}^\infty \) has downward jumps at
PROBABILISTIC AND DETERMINISTIC AVERAGING

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

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

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

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

$$(2.7.2) \quad P(S(L_\tau) \in B) > P(E \cap \{S_{k+j+1} \in B\}) > (\nu^{(k)} \circ \nu^{(j+1)})(B).$$

The convolution $\nu^{(k)} \circ \nu^{(j+1)}$ has support contained in $(0, \infty)$, has positive total mass, and is absolutely continuous since $\nu_\tau$ is. By (2.7.2) the distribution of $S(L_\tau)$ 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 $\alpha$ satisfying the inequalities specified in the theorem. Define strict ladder epochs

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

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

$$(2.8.1) \quad (M_k, X(L_k - 1) + 1, X(L_k - 2), \ldots, X(L_k)), \quad k = 1, 2, \ldots,$$

are i.i.d. Setting $S'_k = S(L_k)$, $k = 0, 1, \ldots$, $X'_k = S'_k - S'_{k-1} = \sum_{t_k-1 < j \leq t_k} X_j$, 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 $E(M'_1) < \infty$, $E((X'_1)^2) < \infty$. 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

$$R_0 = 0, \quad R_k = \sum_{t_k-1 < j \leq t_k} f(S_j), \quad k = 1, 2, \ldots.$$

**Lemma 2.8.1.** $n^{-\alpha} \sum_{k=1}^{n} (R_k - ER_k) \rightarrow 0$, $r$-quickly.

**Proof.** Define the pre-$L_\tau$ $\sigma$-algebra by

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

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

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

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

Next, $|R_k| < MM_k$, so

$$E(|R_k|^p) < ME^{M_k} = ME(M_k^p) < \infty.$$

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

$$(2.8.2) \quad n^{-\alpha} \sum_{k=1}^{n} (R_k - g(S'_{k-1})) \rightarrow 0, \quad r$-quickly.
Now $S'_k = \sum^k_{j=1} X'_j$ where the $X'_j$ are i.i.d. with $E((X'_j)^d) < \infty$, and the d.f. $G$ of $X'_j$ 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 $r < a(l - 1) - 1$ we conclude

$$n^{-a} \sum^n_{i=1} \{ g(S'_i) - Eg(S'_i) \} \to 0, \ r-quickly.$$  

So

$$n^{-a} \sum^n_{i=1} \{ g(S'_{k-1}) - Eg(S'_{k-1}) \}$$

(2.8.3)

$$= \left( \frac{n - 1}{n} \right)^a \cdot (n - 1)^{-a} \cdot n^{-1} \sum^n_{i=1} \{ g(S'_i) - Eg(S'_i) \} \to 0, \ r-quickly.$$  

However, $Eg(S'_{k-1}) = E(E(R_k|S'_{k-1})) = ER_k$. 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

$$n^{-a} \left\{\sum^n_{i=1} f(S_k) - E \sum^n_{i=1} f(S_k) \right\} \to 0, \ r-quickly.$$  

We prove

$$n^{-a} \left\{\sum^n_{i=1} f(S_k) - \sum^{[ne^t]}_{i=1} f(S_k) \right\} \to 0, \ r-quickly,$$

and

$$n^{-a} E \left\{\sum^n_{i=1} f(S_k) - \sum^{[ne^t]}_{i=1} f(S_k) \right\} \to 0.$$  

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

$$n^{-a} \sum^{[ne^t]}_{i=1} \{ f(S_k) -Ef(S_k) \} \to 0, \ r-quickly,$$

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 $n^{-a}M(1 + |L_n - ne^t|)$. Since $L_n = M_1 + \cdots + M_n$ where the $M_k$ are i.i.d. with $EM_1 = e^A$, $E(M_1^d) < \infty$, the Lai strong law [26, (6.5)] yields $n^{-a}(L_n - ne^t) \to 0$ $(\lambda \alpha - 1)$-quickly. But $r < \lambda \alpha - \frac{1}{2} \lambda < \lambda \alpha - 1$, 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.
3. Proofs of Theorems 2 and 2'. Define $S_t = S(t) := S_{[t]}$ for $t > 0$. Let $W(t) = W_t$ be a Wiener process with mean $\mu$ and variance $\sigma^2$. For $T > 0$ define

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

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 $\nu_T$, 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 $\nu$ on the Borel algebra $\mathcal{B}$ in $\mathbb{R}$, define the variation norm by $||\nu|| := \frac{1}{2} \int_{\mathbb{R}} |\nu(\{x\})| \, dx$. The following properties will be needed.

(a) For probability measures $\nu_1, \nu_2$, part of the result known as Scheffé's lemma gives

$$
||\nu_1 - \nu_2|| = \frac{1}{2} \int_{\mathbb{R}} |\nu_1(\{x\}) - \nu_2(\{x\})| = \sup \{\nu_1(B) - \nu_2(B) : B \in \mathcal{B}\}.
$$

The supremum is attained on some $A^+ \in \mathcal{B}$. If $\nu_1, \nu_2$ have densities $h_1, h_2$ then one may identify $A^+$ with (for instance) $\{x : h_1(x) > h_2(x)\}$.

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

(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

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

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

Write $F_t, \Phi_t$ for the d.f.s of $S_t, W_t$ respectively. Proposition 3.1 below establishes the variation-norm closeness of $F_t$ and $\Phi_t$, and Proposition 3.2 deduces that $E\nu_T$ 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 $E(||X||^{2+\delta}) < \infty$ for some $\delta$ with $0 < \delta < 1$. If $\mathcal{G}$ satisfies LLT-I then

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

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

$$
||G_t - \Phi_t|| = O(t^{-\theta/2}), \quad t \to \infty.
$$
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 $n > 1$ write $n = km + r$, $0 < r < m$; then $F_n = (F_m)^{(k)} \ast F_r = (H_m + J_m)^{(k)} \ast F_r$, so that $J_n$, the singular component of $F_n$, is a component of $J_m^{(k)} \ast F_r$, whence

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

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

$$\|F_n - \Phi_n\| \leq \|F_n - \Phi_{\mu, \sigma}^{(n)}\| \leq \|F_n - \Phi_{\mu, \sigma}^{(n)}\| + \|\Phi_{\mu, \sigma}^{(n)} - \Phi_{\mu, \sigma}^{(n)}\|$$

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

Turning to the lattice case, we shall show that

$$\|G_n - \Phi_{\mu, \sigma}^{(n)}\| = O(n^{-\delta/2}), \quad n \to \infty,$$

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 $g_n(x) = P(S_n = k), k < x < k + 1, k \in \mathbb{Z}$. So (3.3) is equivalent to

$$\int_{-\infty}^{\infty} |g_n(x) - \phi_{\mu, \sigma}^{(n)}(x)| \, dx = O(n^{-\delta/2}).$$

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

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

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

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

that is, $\Delta_n := \sup_{x \in \mathbb{R}} \sigma n^{1/2} g_n(x) - \phi_{\mu, \sigma}^{(n)}(x) = O(n^{-\delta/2}).$ 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 $E := \{x: |x - n\mu| < \sigma n^{1/2} \Delta_n^{1 + \theta/\delta}\}$; then

$$\int_{E} |g_n(x) - \phi_{\mu, \sigma}^{(n)}(x)| \, dx \leq 2 \Delta_n^{\theta/\delta} = O(n^{-\theta/2}).$$

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

$$\int_{R \setminus E} \phi_{\mu, \sigma}^{(n)}(x) \, dx = \Phi(-\Delta_n^{-1 + \theta/\delta}) + 1 - \Phi(\Delta_n^{-1 + \theta/\delta})$$

$$= O(n^{-\theta/2}).$$
whence
\[
\int_{\mathbb{R} \setminus E} g_n(x) \, dx = 1 - \int_E g_n(x) \, dx
\]
\[
\leq 1 - \int_E \phi_{n, \sigma^2}(x) \, dx + \int_E \left| g_n(x) - \phi_{n, \sigma^2}(x) \right| \, dx
\]
\[
= O(n^{-\theta/2}).
\]
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 $\varepsilon > 0$,
\[
E(\mu_T - \nu_T) = o(T^{-\delta/2 + \varepsilon}), \quad T \to \infty.
\]

**Proof.** Take $\varepsilon$ 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, $\|F_t - \Phi_t\| < Ct^{-\theta/2}$, all $t > 0$, while under LLT-II, $\|G_t - \Phi_t\| < Ct^{-\varepsilon/2}$, all $t > 0$. Now under LLT-I,
\[
\mathbb{E}(\mu_T - \nu_T) = \left[ \int_0^1 E \left\{ f(S_{Tu}) - f(W_{Tu}) \right\} \right] dt
\]
\[
= \left| \int_0^1 \int_{-\infty}^{\infty} f(x) \left| P(S_{Tu} \in dx) - P(W_{Tu} \in dx) \right| \right| dt
\]
\[
< M \int_0^1 2\|F_{Tu} - \Phi_{Tu}\| \right| dt
\]
\[
< 2M \int_0^1 C(Tu)^{-\theta/2} \, du = O(T^{-\theta/2}) = o(T^{-\delta/2 + \varepsilon}).
\]
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 $\xi$ be uniformly distributed over $[0, 1]$, independent of the process $\{W_t\}_{t \geq 0}$. Observe that $\lambda_T$ may be considered as $E[f(\mu T_\xi)]$ whereas $E\nu_T$ equals $E[f(W(T\xi))]$. 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 $\xi$ as above, let $\mathcal{H}_T, \mathcal{L}_T$ denote the distributions of $\mu T_\xi, W(T\xi)$ respectively. Then for every $\varepsilon > 0$,
\[
\|\mathcal{L}_T - \mathcal{H}_T\| = o(T^{1/2 + \varepsilon}), \quad T \to \infty.
\]

**Proof.** Set $B_t := (W_t - \mu t)/\sigma$, so that $\{B_t\}_{t \geq 0}$ 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
\[
\|\mathcal{L}_T - \mathcal{H}_T\| = \|\mathcal{H}_T - \mathcal{P}\|.
\]
where $\mathcal{N}$ is the uniform distribution on $[0, 1]$ and $\mathcal{N}_T$ is the distribution of $\xi + \sigma(T)\xi B(T\xi)$. Using the scale property of Brownian motion, $\mathcal{N}_T$ is also the distribution of $\xi + B(\xi/T)$, where $T := \mu^2 T/\sigma^2$. Thus $\mathcal{N}_T$ has probability density

\begin{equation}
(3.10) \quad s_T(x) := \int_0^1 (2\pi y/T')^{-1/2} \exp\left\{-\frac{1}{2}(x-y)^2 T'/y\right\} \, dy, \quad x \in \mathbb{R}.
\end{equation}

Fix $x > 0$. The integrand in (3.10), as a function of $y$, defines a probability density on $\mathbb{R}^+$, 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 $\mathcal{N}$ is greater than that of $\mathcal{N}_T$ on the interval $(0, 1)$, and not elsewhere, whence the variation-norm distance evaluates as

\begin{equation}
\| \mathcal{N}_T - \mathcal{N} \| = \int_{\mathbb{R} \setminus (0, 1)} s_T(x) \, dx = P(\xi + B(\xi/T) \notin (0, 1)).
\end{equation}

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

\begin{align*}
P(\xi + B(\xi/T) \notin (0, 1)) &\leq P(\xi < T^{-1/2+\varepsilon/2}) + P(\xi > 1 - T^{-1/2+\varepsilon/2}) \\
&\quad + P(|B(\xi/T)| > T^{-1/2+\varepsilon/2}) \\
&\leq 2T^{-1/2+\varepsilon/2} + P(|B(1/T)| > T^{-1/2+\varepsilon/2})
\end{align*}

(since $|B(1/T)|$ is stochastically larger than $|B(\xi/T)|$)

\begin{align*}
&= 2T^{-1/2+\varepsilon/2} + \Phi(-\mu/\sigma)T^{\varepsilon/2} + 1 - \Phi((\mu/\sigma)T^{\varepsilon/2}).
\end{align*}

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 $2 < l < 3$ we have $2 - \frac{1}{2}l > (l-1)^{-1}$. Hence any value of $\alpha$ 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 $\varepsilon$ in Theorem 2 to be $\alpha - 1 + \frac{1}{2}\delta$. Then Theorem 2 becomes

\begin{equation}
n^{-\alpha}\left\{E\sum_{1}^{n} f(S_k) - \int_{0}^{n} f(\mu x) \, dx\right\} \to 0,
\end{equation}

which with Theorem 1 yields (1.4) as required.

**4. The Poisson case.** Consider the special case where $\mathcal{P}$ is an exponential distribution, of parameter 1 say. Then max$(k: S_k < t)$ 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,

\begin{equation}
n^{-1} \sum_{1}^{n} f(S_k) \to c \quad \text{a.s.} \quad (n \to \infty)
\end{equation}
is equivalent to
\[(4.2) \quad N_t^{-1} \sum_{i=1}^{N(t)} f(S_k) \to c \quad \text{a.s.} \ (t \to \infty),\]
as the successive values of the two left-hand sides coincide. Next,
\[Y_n := \sum_{N(n-1) < k < N(n)} f(S_k) = \sum_{n-1 < S_k < n} f(S_k)\]
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
\[(4.3) \quad n^{-1} \sum_{i=1}^{n} Y_k \to c \quad \text{a.s.} \ (n \to \infty).\]
Recall $M := \sup |f(\cdot)| < \infty$; then $|Y_n| < M(N_n - N_{n-1})$, $\text{var } Y_n < M^2$, $\sum_{i=1}^{\infty} n^{-2} \text{var } Y_n < \infty$. So by a result of Kolomogorov [43, p. 165],
\[n^{-1} \sum_{i=1}^{n} (Y_k - EY_k) \to 0 \quad \text{a.s.} \ (n \to \infty).\]
So $(4.1)$–$(4.3)$ are equivalent to
\[(4.4) \quad n^{-1} \sum_{i=1}^{n} EY_k \to c \quad (n \to \infty).\]
But, conditional on $m$ Poisson points falling in $(n - 1, n]$, these are independently uniformly distributed over $(n - 1, n]$, so
\[EY_n = \sum_{i=1}^{\infty} \left(e^{-1/m!} \cdot m \int_{n-1}^{n} f(y) \, dy\right) = \int_{n-1}^{n} f.\]
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,
\[\text{var } Y_k = \int_{k-1}^{k} f^2, \quad E \left\{ \left( Y_k - \int_{k-1}^{k} f \right)^3 \right\} = \int_{k-1}^{k} f^3.\]
Now $|f_0^{n} f^3| < M f_0^{n} f^2 = o(\{f_0^{n} f^2\}^{3/2})$ so we have Liapounov’s condition for the $Y_k$. By Liapounov’s theorem,
\[\left( \int_0^{n} f^2 \right)^{-1/2} \left( \sum_{i=1}^{n} Y_k - \int_0^{n} f \right) \overset{d}{\to} \Phi \quad (n \to \infty),\]
or
\[(4.5) \quad \left( \int_0^{n} f^2 \right)^{-1/2} \left\{ \sum_{i=1}^{N(n)} f(S_k) - \int_0^{n} f \right\} \overset{d}{\to} \Phi.\]
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
formulation suitable for our case of independent nonidentically distributed summands. Thus
\[ \left( \int_0^{S(n)} f^2 \right)^{-1/2} \left\{ \sum_{i=1}^n f(S_k) - \int_0^{S(n)} f \right\} \to \Phi. \]
If we assume further
\[ (4.6) \quad \liminf_{n \to \infty} n^{-1} \int_0^n f^2 > 0 \]
then \(|f(S(n))f^2/f_0^2 - 1| < Mn^{-1}|S_n - n| \cdot n/f_0^2 \to 0 \text{ a.s.} \); hence we may replace \( f(S(n))f^2 \) by \( f_0^2 \) in the above to obtain
\[ \left( \int_0^n f^2 \right)^{-1/2} \left\{ \sum_{i=1}^n f(S_k) - \int_0^n f \right\} \to \Phi, \]
the CLT for \( F(S_n) \) (note the random centring).

Again, assuming (4.6), Petrov's LIL [43, Corollary 5.2.3] yields
\[ \limsup_{n \to \infty} f^2 f^{2n} f^2 \cdot \log \log f^2 = 1 \quad \text{a.s.}, \]
or
\[ (4.7) \quad \limsup_{n \to \infty} \left\{ 2 \int_0^n f^2 \cdot \log \log f^2 \right\}^{-1/2} \left\{ \sum_{i=1}^{N(n)} f(S_k) - \int_0^n f \right\} = 1 \quad \text{a.s.} \]
Consider for simplicity the important special case where \( f \) is the indicator function \( I_A \) of some fixed set \( A \subset \mathbb{R} \); then \( f^2 = f \). We also assume \( f(\cdot) = \hat{f}(\cdot) \to c \) \((C_1)\), that is, \( |A \cap (0, n)| \sim cn \) as \( n \to \infty \), where \( |\cdot| \) denotes Lebesgue measure. Then (4.6) demands \( c > 0 \). Now
\[ (4.8) \quad n^{-1/2} \left\{ \sum_{i=1}^{N(n)} I_A(S_k) - \sum_{i=1}^n I_A(S_k) \right\} \leq n^{-1/2}|N_n - n|. \]
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
\[ n^{-1/2} \left\{ |A \cap (0, n)| - \sum_{i=1}^n P(S_k \in A) \right\} \]
\[ = n^{-1/2} \left\{ \sum_{i=1}^{N(n)} f(S_k) - \sum_{i=1}^n f(S_k) \right\} = O(1) \]
as in the proof of (2.6.9) above. Combining (4.5), (4.8) and (4.9), we merely obtain the conclusion that
\[ (4.10) \quad n^{-1/2} \sum_{i=1}^n \left\{ I_A(S_k) - P(S_k \in A) \right\}, \quad n = 1, 2, \ldots, \text{ is tight.} \]
Similarly, (4.5) and (4.8) alone yield that
\[ (4.11) \quad n^{-1/2} \left\{ \sum_{i=1}^n I_A(S_k) - |A \cap (0, n)| \right\}, \quad n = 1, 2, \ldots, \text{ is tight.} \]
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)

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

and

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

by the LIL of renewal theory. Hence

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

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

$$\limsup_{n} (2n \log \log n)^{-1/2} \sum_{l=1}^{n} \left\{ I_A(S_k) - P(S_k \in A) \right\} \leq 1 + c^{1/2} \text{ a.s.}$$

5. Complements.

5.1. When strong mixing ('weak $\phi$-mixing') [43, p. 212] is strengthened to $\phi$-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 $\phi$-mixing if and only if there exist constants $c > 0, C$ such that

$$\bar{F}(x + y) \leq Ce^{-\phi} \bar{F}(x), \quad x, y > 0.$$

Then the $\phi$-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 $\mathfrak{F}(x) := \bar{F}(\log x)$ having negative upper Matuszewska index $\alpha\mathfrak{F}$ (Bingham and Goldie [7, III]). 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 $S_n := \sum_{i=1}^{n} X_k$ is not strong mixing (for any i.i.d. random variables $X_k$ of finite variance) consider the events $C_m := \{S_m < m\mu\}$, where $\mu := E X_1$. It is easy to see that as $m \to \infty$ with $n$ fixed,

$$P(C_m \cap C_{m+n}) \to 0;$$

so $\rho_n$ 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
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 Marcin-
ciewicz-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 Marcin-
ciewicz-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 \( \mu^{-1}F(\cdot) \) on \( \mathbb{R}^+ \). Hence

\[
P(B_m < 1, B_{m+n} > n) - P(B_m < 1)P(B_{m+n} > n)
\]

as \( m \to \infty \) with \( n \) fixed. Thus the mixing coefficient \( \rho_n \) in Theorem 4 is at least

\[
c_{n+1} \mu^{-1}F(y) \, dy,
\]

where \( c := \int_0^\infty \mu^{-1}F(x) \, dx > 0 \). If \( E(X_1^+) < \infty \) but \( E(X_1^{1+\epsilon}) = \infty \)
then \( \rho_n = O(n^{-(\ell-1)}) \) but \( \rho_n \neq O(n^{-(\ell-1+\epsilon)}) \). 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 \( \rho_n \).

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

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

\[
\|\mathcal{C}_T - \mathcal{K}_T\| = o(T^{-1/2+\epsilon}), \quad T \to \infty.
\]

(Define variation norms as in \( \mathbb{R} \), by \( \|\nu\| := \frac{1}{2} \int_{\mathbb{R}^k} |\nu|(dx) \), for a signed measure \( \nu \)
on the Borel sets of \( \mathbb{R}^k \).) 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 \((S_n \text{ integer valued})\), suppose one has a sequence of uniformly
bounded random variables \( Z_1, Z_2, \ldots, \) independent of each other and of
\( \{S_n\}_{n \geq 0} \). Then if \( \mathcal{P} \) satisfies the conditions of Theorem 2' we can obtain, for \( r, \alpha \) 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 \( \mathbb{R} \), 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
(the LLT-II case) we set \( m(x) := m(k) \) for \( k < x < k + 1 \). 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 \( \mathcal{D}_x \), for \( k = 1, 2, \ldots \). Theorems 1, 2 and 2' remain in force, their conclusions (1.2), (1.3) and (1.4) becoming respectively

\[
(1.2') \quad n^{-a} \sum_{k=1}^{n} \{ V_k - EV_k \} \to 0, \quad r\text{-quickly},
\]

\[
(1.3') \quad n^{-1} \sum_{k=1}^{n} EV_k - n^{-1} \int_0^n m(\mu x) \, dx = o(n^{-\delta/2+\epsilon}), \quad n \to \infty,
\]

\[
(1.4') \quad n^{-a} \left\{ \sum_{k=1}^{n} V_k - \int_0^n m(\mu x) \, dx \right\} \to 0, \quad r\text{-quickly}.
\]

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 \( \mathcal{D}_x \) distribution in §3. Thus define

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

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 \{\mathcal{W}_i\}_{i>0} \( \tilde{V}_t \) has distribution \( \mathcal{D}_{\mathcal{W}_t} \), 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 \( n^{-1} \sum_{k=1}^{n} f(S_k) \) must be degenerate. By the usual subsequence argument, any in-probability limit of \( n^{-1} \sum_{k=1}^{n} f(S_k) \) 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 \( p = 0 \) 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 \( (2x)^{-1} \int_x^\infty f(y) \, dy \to c (x \to \infty) \);

(b) [12] if \( x^{-1} \int_0^x f(y) \, dy \to p, \quad x^{-1} \int_0^x f(y) \, dy \to q (x \to \infty) \), then \( n^{-1} \sum_{0}^{n} f(S_k) \) 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 \( n^{-1} \sum_{0}^{n} f(S_k) \) converges in law, then

\[
\lim_{x \to \infty} x^{-1} \int_0^x f(y) \, dy \text{ exists}.
\]

Note that when \( f \) vanishes on \(( -\infty, 0 ) \), Theorem B gives the equivalence of \( f(x) \to p (C_1), x \to \infty \),

\[
Ef(S_n) \to \frac{1}{2} p, \quad n \to \infty,
\]

\[
n^{-1} \sum_{0}^{n} f(S_k) \text{ converges in law},
\]

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
$\mu > 0$ 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:

1. $Ef(S_n) \to c$ (as $n \to \infty$);
2. $(\epsilon \sqrt{x})^{-1}\int_x^{\infty} f(y) \, dy \to c \ (x \to \infty)$ for some (all) $\epsilon > 0$;
3. $\int_0^x f(y) \, dy = c + o(x^{-1/2}) + \int_0^x \epsilon(y) \, dy, \, \epsilon(x) \to 0 \ as \ x \to \infty$;
4. $f(n) \to c \ (B) \ [lattice \ case \ only]$;
5. $f(n) \to c \ (E_p) \ for \ some \ (all) \ p \in (0, 1) \ [lattice \ case \ only]$.

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], Kulinić [25], Taraskin [46]. Baxter and Brosamler [2] consider a.s. limits of $\int_0^X f(X_s) \, ds$ for diffusions $X$ on compact metric spaces.

For *integrable* $f$, the a.s. convergence of $n^{-1}\Sigma^n_0 f(S_k)$ 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 $n^{-1}\Sigma^n_1 f(S_k)$ $\to c$ a.s. to $\int_0^n f(\mu) \, dy \to c$. 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 $n^{-1}\Sigma^n_1 f(S_k) \sim c$ for large $n$ (cf. e.g. Niederreiter [31]). More generally, for any bounded measurable $f$, Theorem 2' shows that the value of $\int_0^n f(\mu) \, dt$ can be approximated with error $o(n^a)$ 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, $\sum_{k=0}^{n} f(k + U_k)$ ($U_k$ independent and uniform on $[0, 1]$) is easily seen to have as good convergence properties as $\sum_{k=0}^{n} f(S_k)$ (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 $D \subset \mathbb{N}$ the set of defectives, one can sample at the instants of a random walk $\{S_n\}$, and then $n^{-1} \sum_{k=0}^{n} I_D(S_k)$ is a strongly consistent estimator for $p$, with error $o(n^{1-a})$.

References

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


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