ASYMPTOTIC EXPANSIONS IN MULTIDIMENSIONAL MARKOV RENEWAL THEORY AND FIRST PASSAGE TIMES FOR MARKOV RANDOM WALKS

by

Cheng-Der Fuh
Tze Leung Lai

Technical Report No. 2001-17
May 2001

Department of Statistics
STANFORD UNIVERSITY
Stanford, California 94305
ASYMPTOTIC EXPANSIONS IN MULTIDIMENSIONAL MARKOV RENEWAL
THEORY AND FIRST PASSAGE TIMES FOR MARKOV RANDOM WALKS

by

Cheng-Der Fuh
Institute of Statistical Science
Academia Sinica, Taipei, Taiwan

Tze Leung Lai
Department of Statistics
Stanford University

Technical Report No. 2001-17
May 2001

This research was supported in part by National Science Foundation
grant DMS-00-72523 and by National Security Agency grant
MDA 904-00-1-0018

Department of Statistics
Sequoia Hall
STANFORD UNIVERSITY
Stanford, California 94305

http://www-stat.stanford.edu
ASYMPTOTIC EXPANSIONS IN MULTIDIMENSIONAL MARKOV RENEWAL THEORY AND FIRST PASSAGE TIMES FOR MARKOV RANDOM WALKS

CHENG-DER FUH* AND TZE LEUNG LAI**

Academia Sinica and Stanford University

We prove a $d$-dimensional renewal theorem, with an estimate on the rate of convergence, for Markov random walks. This result is applied to a variety of boundary crossing problems for a Markov random walk $\{(X_n, S_n), n \geq 0\}$, in which $X_n$ takes values in a general state space and $S_n$ takes values in $\mathbb{R}^d$. In particular, for the case $d = 1$, we use this result to derive an asymptotic formula for the variance of the first passage time when $S_n$ exceeds a high threshold $b$, generalizing Smith's classical formula in the case of i.i.d. positive increments for $S_n$. For $d > 1$, we apply this result to derive an asymptotic expansion of the distribution of $(X_T, S_T)$, where $T = \inf\{n : S_{n,1} > b\}$ and $S_{n,1}$ denotes the first component of $S_n$.

Abbreviated title: Markov renewal theory and first passage problems

AMS 2000 subject classifications. Primary 60K05; secondary 60J10, 60K15.

Key words and phrases. First passage time; Markov random walk; Markov renewal theory; rate of convergence.

*Postal address: Institute of Statistical Science, Academia Sinica, Taipei 11529, Taiwan, ROC.
**Postal address: Department of Statistics, Sequoia Hall, Stanford University, Stanford, CA 94305-4065, USA.
1. Introduction

Let \( \{X_n, n \geq 0\} \) be a Markov chain on a general state space \( \mathcal{X} \) with \( \sigma \)-algebra \( \mathcal{A} \) and let \( S_n = \sum_{k=0}^{n} \xi_k \) for \( n \geq 0 \) be the additive component of a Markov random walk \( \{(X_n, S_n), n \geq 0\} \), which is a Markov chain on \( \mathcal{X} \times \mathbb{R}^d \) such that

\[
P\{(X_1, S_1) \in A \times (B + s) \mid (X_0, S_0) = (x, s)\} = P\{(X_1, S_1) \in A \times B \mid (X_0, S_0) = (x, 0)\} = P(x, A \times B),
\]
for all \( x \in \mathcal{X}, \ s \in \mathbb{R}^d, A \in \mathcal{A} \) and \( B \in \mathcal{B} \) (=: Borel \( \sigma \)-algebra on \( \mathbb{R}^d \)). For an initial distribution \( \nu \) on \( (X_0, S_0) \), let \( P_\nu \) denote the probability measure under the initial distribution \( \nu \) on \( (X_0, S_0) \) and let \( E_\nu \) denote expectation under the initial distribution \( \nu \) on \( (X_0, S_0) \). In the case where \( \nu \) is degenerate at \( (x, 0) \), we denote \( P_\nu \) by \( P_x \) and \( E_\nu \) by \( E_x \). Suppose the Markov chain \( X_n \) has stationary distribution \( \pi \). For the case \( d = 1 \), let \( g : \mathcal{X} \times \mathbb{R} \to \mathbb{R} \). The classical Markov renewal theorem [cf. Smith (1958), Cinlar (1969), Kesten (1974), Athreya, McDonald and Ney (1978), Shurenkov (1984) and Alsmeyer (1994)] states that under certain regularity conditions,

\[
E_\nu \left( \sum_{k=0}^{\infty} g(X_k, b - S_k) \right) \to \int_{\mathcal{X}} \int_{\mathbb{R}} g(x, s)dsd\pi(x)/\int_{\mathcal{X}} E_x \xi_1 d\pi(x)
\]
as \( b \to \infty \). In Theorem 4 we establish rates of convergences for (1.2), generalizing Stone’s (1965) results in the i.i.d. case.

Multidimensional renewal theory (for the case \( d > 1 \)) with convergence rates has been developed in the case where the increments \( \xi_k \) of the random walk are i.i.d. by Stam (1968, 1971), Carlsson (1982), Carlsson and Wainger (1982) and Keener (1988, 1990). Theorems 1-3 generalize these results to Markov random walks. Our approach uses the Fourier transform of the Markov transition operator and Schwartz’s theory of distributions. This approach and the proofs of Theorems 1-4 are given in Section 4.

After reviewing some basic definitions and concepts, Section 2 states these Markov renewal theorems. Markov renewal theory is often applied to the ladder (instead of the original) random walk. Examples are given in Section 2 to show how Theorems 1-4 can be applied to the ladder variables. Section 3 describes some applications of these theorems to a variety of first passage problems for Markov random walks.

2. Markov renewal theorems with convergence rates

Let \( \{(X_n, S_n), n \geq 0\} \) be a Markov random walk on \( \mathcal{X} \times \mathbb{R}^d \). Let \( w : \mathcal{X} \to [1, \infty) \) be a measurable function and let \( \mathcal{B} \) be the Banach space of measurable functions \( h : \mathcal{X} \to \mathcal{C} \) (:= set of complex numbers) with \( ||h||_w := \sup_x |h(x)|/w(x) < \infty \). The Markov chain \( \{X_n, n \geq 0\} \)
is assumed to be irreducible (with respect to some measure on $A$), aperiodic and $w$-uniformly ergodic, i.e., there exists an invariant probability measure $\pi$ such that $\int w(y)d\pi(y) < \infty$ and

\begin{align}
\lim_{n \to \infty} \sup_x \{|E_x(h(X_n)) - \int h(y)d\pi(y)/w(x) : x \in X, |h| \leq w\} &= 0, \\
\sup_x \{E_x(w(X_1))/w(x)\} &= \infty,
\end{align}

(2.1) \hspace{1cm} (2.2)

cf. Chapter 16 of Meyn and Tweedie (1993). In addition, it will be assumed that for some $r \geq 2$,

\begin{equation}
\sup_x \{E_x(|\xi_1|^r w(X_1))/w(x)\} < \infty.
\end{equation}

(2.3)

Under irreducibility and aperiodicity, condition (2.1) implies that there exist $C > 0$ and $0 < \rho < 1$ such that for all $h \in B$ and $n \geq 1$,

\begin{equation}
\sup_x |E_x(h(X_n)) - \int h(y)d\pi(y)/w(x)| \leq C\rho^n \|h\|_w;
\end{equation}

see pages 382-383 of Meyn and Tweedie (1993).

When the increments $\xi_n$ are i.i.d. and strongly nonlattice, Stone (1965) and Carlsson (1983) derived the rate of convergence of the renewal measure to its limit under moment conditions on $\xi_n$ in the case $d = 1$, while Carlsson and Wainger (1982) and Keener (1990) developed asymptotic expansions of the renewal measure in the case $d > 1$. To generalize these results to Markov random walks, we first extend Cramér’s condition (corresponding to strongly nonlattice random vectors): Let $\pi$ be the stationary distribution of $\{X_n, n \geq 0\}$ and let $P_\pi$ denote $\int P_\pi d\pi(x)$ and $E_\pi$ be expectation under $P_\pi$. The Markov random walk $\{(X_n, S_n), n \geq 0\}$ is called strongly nonlattice if

\begin{equation}
\limsup_{|\theta| \to \infty} |E_\pi e^{i\theta' \xi_1}| < 1.
\end{equation}

(2.5)

It is called nonlattice if $|E_\pi e^{i\theta' \xi_1}| < 1$ for every $\theta \neq 0$. Here and in the sequel, we use column vectors to denote $\theta \in \mathbb{R}^d$, $\theta'$ to denote the transpose of $\theta$, and $|\theta|$ to denote its Euclidean norm $(\theta' \theta)^{1/2}$.

Let $\mu = E_\pi \xi_1$ and $V = \lim_{n \to \infty} n^{-1} E_\pi \{(S_n - n\mu)(S_n - n\mu)\}'$, which are well defined under (2.1) and (2.3). Let $S_{n,j}$ (or $\xi_{n,j}$, $\mu_j$, $\theta_j$) denote the jth component of the d-dimensional vector $S_n$ (or $\xi_n$, $\mu$, $\theta$). Suppose $\mu_1 > 0$. Without loss of generality, it will be assumed that $V$ is positive definite (i.e., $\xi_n$ is strictly d-dimensional under $\pi$), because otherwise we can consider a lower-dimensional subspace instead. In the case $d > 1$ define

\begin{equation}
\gamma = E_\pi \{(\xi_{n,2}/\mu_1, \cdots, \xi_{n,d}/\mu_1)'\}, \quad \tilde{V} = (-\gamma, I_{d-1})V\begin{pmatrix} -\gamma' \\ I_{d-1} \end{pmatrix},
\end{equation}

(2.6)

where $I_k$ is the $k \times k$ identity matrix. Note that $\tilde{V}$ is the asymptotic covariance matrix (under $P_\pi$) of $\{(S_{n,2}, \cdots, S_{n,d})' - S_{n,1}\gamma\}/\sqrt{n}$. For $s \in \mathbb{R}^d$, define $\ddot{s} = (s_2, \cdots, s_d)' - s_1 \gamma$. 

2
First consider the case of i.i.d. $\xi_n$, with $d > 1$ and $S_0 = 0$. The renewal measure is defined by $U(B) = \sum_{n=0}^{\infty} P\{S_n \in B\}$, and multivariate renewal theory is concerned with approximating $U(s + \cdot)$ by $\Psi_k(s + \cdot)$ as $s_1 \to \infty$, where $\Psi_k$ is a $\sigma$-finite measure on $\mathbb{R}^d$ whose density function (i.e., Radon-Nikodym derivative) with respect to Lebesgue measure is of the form

$$\psi_k(s) = \frac{1}{\mu_1 \sqrt{\det V}} \left( \frac{\mu_1}{2 \pi s_1} \right)^{(d-1)/2} e^{-\mu_1 s_1^\gamma V^{-1} \bar{s}^2/2s_1} \left( 1 + \sum_{j=1}^{k} s_1^{-j/2} \omega_j(\bar{s}/\sqrt{s_1}) \right)$$

for $s_1 > 0$, and $\psi_k(s) = 0$ for $s_1 \leq 0$, where $\omega_j(u) = \sum_{l=0}^{\infty} q_l(u)$ and $q_l(u)$ is a polynomial of degree $l$ in $u$ whose coefficients are associated with the Taylor expansion of $(1 - \text{e}^{i \theta \xi_1})^{-1}$ near $\theta = 0$; see Carlsson (1982), Carlsson and Wainger (1982) and Keener (1990). For Markov random walks, the renewal measure involves not only $\{S_n\}$ but also $\{X_n\}$. For $A \in A$ and $B \in B$, define

$$U^A_\nu(B) = \sum_{n=0}^{\infty} P_\nu\{X_n \in A, S_n \in B\}.$$  

We can approximate $U^A_\nu(s + \cdot)$ by $\pi(A)\Psi^A_\nu(s + \cdot)$, in which $\Psi^A_\nu$ is a $\sigma$-finite measure on $\mathbb{R}^d$ with density function $\psi^A_\nu$ with respect to Lebesgue measure, where $\psi^A_\nu(s) = 0$ for $s_1 \leq 0$ and $\psi^A_\nu(s)$ is given by (2.7) for $s_1 > 0$, with the coefficients of the polynomials $\omega_1(\bar{s}), \ldots, \omega_k(\bar{s})$ depending also on $A$ and $\nu$ via Taylor's expansion of the Fourier transform of $U^A_\nu$ near the origin, assuming that

$$E_\nu\{w(X_1)(1 + |S_1|^r)\} < \infty$$

for some sufficiently large $r$ (depending on $k$). Note that when $\nu$ is degenerate at $(x, 0)$, (2.9) follows from (2.3). The precise definition of $\omega_j$ is given in Section 4.1, where we also prove the following multidimensional Markov renewal theorem with bounds on the remainders in approximating $U^A_\nu(s + \cdot)$ by $\pi(A)\Psi^A_\nu(s + \cdot)$ as $s_1 \to \infty$, recalling the assumption $\mu_1 > 0$.

**Theorem 1.** Let $k \geq 1$ and let $\{(X_n, S_n), n \geq 0\}$ be a strongly nonlattice Markov random walk satisfying (2.1)-(2.3) and (2.9) for some $r > k + 5 + \max\{1, (d - 1)/2\}$. Let $A \in A$ and $B$ be a $d$-dimensional rectangle $\prod_{j=1}^{d} [\alpha_j, \beta_j]$. Then as $s_1 \to \infty$,

$$U^A_\nu\left(s + \begin{pmatrix} 0 \\ s_1 \gamma \end{pmatrix} + B\right) = \pi(A)\Psi^A_\nu\left(s + \begin{pmatrix} 0 \\ s_1 \gamma \end{pmatrix} + B\right) + o(s_1^{-(d-1+k)/2})$$

uniformly in $\bar{s}$.

**Theorem 2.** Let $\{(X_n, S_n), n \geq 0\}$ be a strongly nonlattice Markov random walk satisfying (2.1)-(2.3) and (2.9) for some $r > 3$. Let $h > 0$ and $\alpha > 0$. Let $B_\alpha$ be the class of all Borel subsets of $\mathbb{R}^{d-1}$ such that $\int_{\partial B} \exp(-|y|^2/2)dy = O(e^\alpha)$ as $\varepsilon \downarrow 0$, where $\partial B$ denotes the boundary of $B$ and $(\partial B)^\varepsilon$ denotes its $\varepsilon$-neighborhood. Then as $s_1 \to \infty$,

$$U^A_\nu([s_1, s_1 + h] \times \sqrt{s_1}(s_1 \gamma + C)) = \pi(A)\Psi^A_\nu([s_1, s_1 + h] \times \sqrt{s_1}(s_1 \gamma + C)) + o(s_1^{-(1+\delta)/2})$$
for every \( \delta < \min(1, r - 3) \), uniformly in \( A \in A \) and \( C \in B_\alpha \).

Theorem 1 is an extension of a corresponding result of Carlsson and Wainger (1982) for i.i.d. \( \xi_n \), while Theorem 2 (which deals with a much wider class of sets \( B \) than that considered in Theorem 1) is an extension of a result of Keener (1990) who derived it as a corollary of his Theorem 3. For \( \varepsilon > 0 \) and \( f : \mathbb{R}^d \to \mathbb{R} \), define the oscillation function \( \Omega_f(s; \varepsilon) = \sup\{|f(s) - f(t)| : |s - t| \leq \varepsilon\} \). Let \( \mathcal{F}_b \) be the set of all Borel functions \( f : \mathbb{R}^d \to [0, 1] \) such that \( f(s) = 0 \) whenever \( s \not\in [b, b + h] \), with fixed \( h > 0 \). We extend Keener's (1990) Theorem 3 to Markov random walks in the following.

**Theorem 3.** Let \( \{(X_n, S_n), n \geq 0\} \) be a strongly nonlattice Markov random walk satisfying (2.1)-(2.3) and (2.9) for some \( r > 3 \). Let \( 0 < \delta < \min(1, r - 3) \). Then for every \( \eta > 0 \), as \( b \to \infty \),

\[
\int f(s) dU^A_\eta(s) = \pi(A) \int f(s) d\Psi^{A, \nu}_1(s) + O\left(\int \Omega_f(s; b^{-\eta}) d\Psi^{A, \nu}_1(s)\right) + o(b^{-(1+\delta)/2})
\]

uniformly in \( f \in \mathcal{F}_b \) and \( A \in A \).

In the case \( d = 1 \), \( V \) is a scalar, which will be denoted by \( \sigma^2 \). The following theorem provides bounds on the difference between \( U^A_\eta([b, b + h]) \) and its renewal-theoretic approximation as \( b \to \infty \), extending the results of Stone (1965) to Markov random walks.

**Theorem 4.** Suppose \( d = 1 \) and \( \{(X_n, S_n), n \geq 0\} \) is a strongly nonlattice Markov random walk satisfying (2.1)-(2.3) and (2.9) for some \( r \geq 2 \). Then as \( b \to \infty \),

\[
U^A_\eta([b, b + h]) = \pi(A) h/\mu + o(b^{-(r-1)})
\]

uniformly in \( A \in A \).

In the remainder of this section, we give examples of strongly nonlattice Markov random walks satisfying (2.1)-(2.3) and (2.9) for some weight function \( w \), and such that the underlying Markov chain \( \{X_n, n \geq 0\} \) is irreducible and aperiodic. Many time series and queuing models \( X_n \) are irreducible, aperiodic and \( w \)-uniformly ergodic Markov chains, as shown in Chapters 15 and 16 of Meyn and Tweedie (1993), and conditions (2.3) and (2.9) are weak moment conditions on the additive components attached to \( X_n \) that are satisfied in typical applications. However, renewal theorems are often applied to the ladder random walk, as in Section 3 below. The techniques used by Meyn and Tweedie (1993) to prove the \( w \)-uniform ergodicity of a rich class of time series and queuing models can also be applied to show that their ladder random walks indeed satisfy conditions (2.1)-(2.3) and (2.9), as illustrated by the following examples. Recall
that the (positive) ladder epoch of a Markov random walk \( \{(X_n, S_n), n \geq 0\} \) taking values in \( \mathcal{X} \times \mathbb{R} \) is defined by
\[
\tau_+ = \inf\{n \geq 1 : S_n > 0\}.
\]

For \( A \in \mathcal{A} \) and Borel subset \( B \) of \( (0, \infty) \), define
\[
P_+(x, A \times B) = P\{X_{\tau_+} \in A, S_{\tau_+} \in B | X_0 = x\}.
\]

The kernel \( P_+ \) is the transition probability kernel of a Markov random walk that has the ladder chain as the underlying Markov chain on \( \mathcal{X} \). The following examples show how \( w \)-uniform ergodicity of the ladder chain and finiteness of moments of \( S_{\tau_+} \) can be established.

**Example 1.** Meyn and Tweedie (1993), pages 380 and 383, establish \( w \)-uniform ergodicity of the AR(1) model \( X_n = \alpha X_{n-1} + Z_n \) by proving that a drift condition is satisfied, where \( |\alpha| < 1 \) and the \( Z_n \) are i.i.d. random variables, with \( E|Z_n| < \infty \), whose common density function \( q \) with respect to Lebesgue measure is positive everywhere. Suppose the conditional distribution of \( \xi_n \) given \( X_0, \ldots, X_n \) is of the form \( F_{X_{n-1}, X_n} \) such that
\[
\limsup_{|\theta| \to 0} \int_{\mathcal{X}} \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} e^{i\theta \xi} dF_{x, \alpha x + z}(\xi) \right\} q(z) dzd\pi(x) < 1,
\]
where \( \pi \) is the stationary distribution of \( \{X_n\} \). Let \( S_n = \sum_{i=1}^{n} \xi_i \), \( S_0 = 0 \). Then \( \{(X_n, S_n), n \geq 0\} \) is a strongly nonlattice, and by Theorem 6 (ii) of Fuh and Lai (1998), so is the ladder random walk with transition kernel \( P_+ \) defined by (2.11). Assume furthermore that
\[
\sup_x E_\pi(\xi_1|1 + |Z_1|) < \infty \quad \text{and} \quad \mu := E_\pi \xi_1 > 0.
\]

Let \( w(x) = |x| + 1 \). To show that the ladder chain is \( w \)-uniformly ergodic, first note that \( X_{\tau_+} \) has a positive density function with respect to Lebesgue measure \( \lambda \) and is therefore \( \lambda \)-irreducible. Moreover, the ladder chain is clearly aperiodic; see Section 5.4.3 of Meyn and Tweedie (1993). Hence by Theorem 16.0.1 of Meyn and Tweedie (1993), it suffices to show that there exist positive constants \( b, \beta \) and a petite set \( C \) such that
\[
E_\pi w(X_T) - w(x) \leq -\beta w(x) + b1_C(x) \quad \text{for all } x \in \mathbb{R},
\]
for \( T = \tau_+ \). We first show that the drift condition (2.14) in fact holds for all stopping times \( T \) (with respect to the filtration generated by \( (X_n, S_n) \)) such that for some \( a > 0 \),
\[
E_\pi T \leq a(|x| + 1) \quad \text{for all } x \in \mathbb{R}.
\]
We then show that \( \tau_+ \) satisfies (2.15) and therefore (2.14) indeed holds for \( T = \tau_+ \).

Let \( Z_1^* = Z_11_{\{|Z_1| \leq B\}}, Z_1^{**} = Z_11_{\{|Z_1| > B\}} \), and note that
\[
|X_T| = \alpha^T X_0 + \alpha^{T-1} Z_1 + \ldots + \alpha Z_{T-1} + Z_T \leq |\alpha||x| + (1 - |\alpha|)^{-1}B + \sum_{i=1}^{T} |Z_i^{**}| \quad \text{for } X_0 = x.
\]
Since the $Z_{i^*}$ are i.i.d., Wald’s lemma yields

\[(2.17) \quad E_x \left| \sum_{i=1}^{T} Z_{i^*} \right| = (E_x T) E|Z_{1^*}| \leq a(|x| + 1) E|Z_{1^*}|,\]

by (2.15). Since $E|Z_{1}| < \infty$, we can choose $B$ sufficiently large so that $aE|Z_{1^*}| + |\alpha| < 1$. In view of (2.16) and (2.17), we can then choose $0 < \beta < 1 - |\alpha| - aE|Z_{1^*}|$, $b > (1 - |\alpha|)^{-1}B + aE|Z_{1^*}|$ and $C = \{ z : |z| \leq K \}$ with $K$ sufficiently large such that (2.14) holds. Note that $C$ is a petite set; see Section 5.2 of Meyn and Tweedie (1993).

To show that $\tau_+$ satisfies (2.15), we use Wald’s lemma for Markov random walks: For any stopping time $T$ with $E_\nu T < \infty$ and $E_\nu w(X_T) < \infty$,

\[(2.18) \quad E_\nu S_T = \mu E_\nu T + E_\nu \{ g(X_T) - g(X_0) \},\]

where $\sup_x |g(x)|/w(x) < \infty$; see Fuh and Zhang (2000) and Section 3.3 below for the definition of $g$. Let $\xi^{(B)}_1 = \xi_1 1\{\xi_1 \leq B\}$, $S_n^{(B)} = \xi^{(B)}_1 + \cdots + \xi^{(B)}_n$ and $\tau(B) = \inf \{ n : S_n^{(B)} > 0 \}$. Since $\mu = E_\nu \xi_1 > 0$, we can choose $B$ large enough so that $\mu^{(B)} = E_\nu \xi_1^{(B)} > 0$. Note that $S_n^{(B)} \leq S_n$ and $\tau(B) \geq \tau_+$. Hence it suffices to show that $\tau(B)$ satisfies (2.15). By the monotone convergence theorem, we need only show that (2.15) holds with $T = \tau(B)$ and $m$ for every $m \geq 1$. Since $S_{\tau(B)\wedge m} \leq B$, (2.18) yields

\[(2.19) \quad B \geq \mu^{(B)}E_x T - E_x |g(X_T)| - |g(x)| \geq \mu^{(B)}E_x T - cE_x |X_T| - c(|x| + 2),\]

since $|g(x)| \leq cw(x) = c(|x| + 1)$ for some $c > 0$ and all $x$. By (2.16) and (2.17),

\[(2.20) \quad E_x |X_T| \leq |\alpha||x| + (1 - |\alpha|)^{-1}B + (E_x T) E(|Z_1| 1_{\{|Z_1| > B\}}).\]

Choosing $B$ large enough so that $cE|Z_{1^*}| 1_{\{|Z_{1^*}| > B\}} < \mu^{(B)}/2$, we obtain from (2.19) and (2.20) that

\[B + c(1 - |\alpha|)^{-1}B + 2c + c|x|(1 + |\alpha|) \geq \mu^{(B)}E_x T/2,\]

proving (2.15) for $T = \tau(B) \wedge m$.

From (2.16) and (2.17) with $T = \tau_+$, it follows that $\sup_x \{ E_x w(X_{\tau_+})/w(x) \} < \infty$. Hence (2.2) holds for the ladder chain. To prove that (2.3) holds for the kernel (2.11), we assume the additional moment conditions

\[(2.21) \quad \sup_x E_x (\xi_1^{\tau_+})^p < \infty \quad \text{and} \quad E|Z_{1}|^{p/(p-1)} < \infty \quad \text{for some} \quad p > 1.\]

First note that $E_x S_{\tau_+} w(X_{\tau_+}) \leq E_x \xi_1^{\tau_+} w(X_{\tau_+}) \leq \{ E_x \xi_1^{\tau_+} \}^{1/p} \{ E_x w^q(X_{\tau_+}) \}^{1/q}$, where $q^{-1} = 1 - p^{-1}$. Let $T = \tau_+$. Then

\[(2.22) \quad E_x \xi_1^{\tau_+} \leq E_x \{ \sum_{i=1}^{T} (\xi_i^{\tau_+})^p \} = E_x \{ \sum_{i=1}^{T} E[(\xi_i^{\tau_+})^p |\mathcal{F}_{i-1}] \} \leq \sup_y E_y (\xi_i^{\tau_+})^p E_x T.\]
From (2.16), there exists a constant \( c_q \) depending only on \( q \) such that

\[
E_x |X_T|^q \leq c_q \{|x|^q + (1 - |\alpha|^{-1})E_x \max_{i \leq T} |Z_i|^q \}
\]

\[
\leq c_q \{|x|^q + (1 - |\alpha|^{-1})E_x \sum_{i=1}^T |Z_i|^q \} = c_q \{|x|^q + (1 - |\alpha|^{-1})E |Z_1|^q E_x T \}.
\]

From (2.2), (2.23) and (2.15), it follows that \( \sup_x E_x (S_{T, \nu}(w(x)))/w(x) < \infty \). Since \( S_0 = 0 \) and \( \sup_x E_x S_{T, \nu}^r/w(x) < \infty \), (2.9) also holds for the kernel \( P_+ \) if the initial distribution \( \nu \) satisfies \( \int_{-\infty}^{\infty} |x|d\nu(x) < \infty \). Hence (2.1)-(2.3) and (2.9) are satisfied by the ladder random walk with transition kernel \( P_+ \) when the underlying chain is an AR(1) model and \( \{\xi_n\} \) satisfies (2.12), (2.13) and (2.21).

**Example 2.** The preceding argument can be easily extended to the case where the underlying chain is a \( k \times 1 \) vector defined by the stochastic difference equation \( X_n = AX_{n-1} + Z_n \), in which \( A \) is a \( k \times k \) matrix with eigenvalues \( \lambda_1, \ldots, \lambda_k \) (possibly complex and not necessarily distinct) such that \( \max_{1 \leq i \leq k} |\lambda_i| < 1 \). We use here \( |\cdot| \) to denote the modules of a complex number or the Euclidean norm of a vector in \( \mathbb{R}^d \). We can further generalize the preceding argument (for \( d = 1 \)) to the case \( d > 1 \), assuming \( \mu_1 > 0 \). Let \( \tau_+ = \inf\{n \geq 1 : S_n, \nu > 0\} \) and generalize (2.11) to the form \( P_+(x, A \times B) = P_x (X_{\tau_+} \in A, S_{\tau_+} \in B) \) for Borel subsets \( A \) of \( \mathbb{R}^k \) and \( B \) of \( \mathbb{R}^d \). With \( w(x) = |x| + 1 \) as before, we can use the same ideas as those in Example 1 to show that (2.1)-(2.3) and (2.9) hold for the kernel \( P_+ \) under the assumptions (2.12) (with \( \theta \xi \) replaced by \( \theta' \xi \)), (2.13) (with \( \mu \) replaced by \( \mu_1 \)), (2.21) and \( \int |x|d\nu(x) < \infty \). Note in this connection that Theorem 6 of Fuh and Lai (1998) can be generalized to multivariate \( S_n \) by making use of "dual pairs" of stopping times introduced by Greenwood and Shaked (1978) to establish Wiener-Hopf factorization for \( d \)-dimensional random walks.

**Example 3.** The conclusions in Example 2 can be further extended to nonlinear stochastic difference equations \( X_n = f(X_{n-1})X_{n-1} + g(X_{n-1})Z_{n-1} \) where \( f(\cdot) \), \( g(\cdot) \) are nonrandom \( k \times k \) matrices such that

\[
\sup_x ||f(x)|| < 1 \quad \text{and} \quad \sup_x ||g(x)|| < \infty,
\]

in which \( ||A|| = \sup_{|x|=1} ||Ax|| \) is the norm of a matrix \( A \). This includes a large class of nonlinear time series models; see Chen and Tsay (1993) whose appendix shows how the first condition in (2.24) can be used.

**Example 4.** Let \( r \geq 1 \) and let \( \{X_n, n \geq 0\} \) be an irreducible Markov chain on a finite state space \( \mathcal{X} \). Let \( \pi \) be the stationary distribution of \( \{X_n, n \geq 0\} \). Suppose the Markov random walk \( \{(X_n, S_n), n \geq 0\} \) is strongly nonlattice, with \( E_\pi \xi_{1,1} > 0 \) and \( E_x |\xi_1|^{r} < \infty \) for every \( x \in \mathcal{X} \), and such that the kernel \( P_+(x, A \times \mathbb{R}^d) \) is aperiodic. Then \( X_n := X_{\tau_n} \) has stationary distribution \( \pi_+ \), where \( \tau_n \) is the \( n \)-th ascending ladder epoch. Combining this with Theorems 3 and 6(ii) of Fuh and Lai (1998) shows that (2.1)-(2.3) hold for the ladder random walk.
\{(\hat{X}_n, \hat{S}_n), n \geq 0\} \text{ with } w \equiv 1 \text{ and } X_n, S_n, \pi \text{ replaced by } \hat{X}_n, \hat{S}_n, \pi_+, \text{ where } \hat{X}_n = X_{\tau_n} \text{ and } \hat{S}_n = S_{\tau_n}.

3. Limit theorems for first passage times of Markov random walks

In this section, we generalize several limit theorems for first passage times of random walks with i.i.d. increments \(\xi_n\) to Markov random walks by making use of Markov renewal theory for the ladder random walk, with kernel (2.11) in which \(\tau_+\) is defined by (2.10) in the case \(d = 1\) and by \(\tau_+ = \inf\{n \geq 1 : S_{n,1} > 0\}\) in the case \(d > 1\). \textit{It is assumed throughout this section that } \(P_x(\tau_+ < \infty) = 1\) \textit{for all } \(x \in \mathcal{X}\) \textit{and that the ladder random walk is strongly nonlattice and satisfies conditions (2.1)-(2.3). Let } \pi_+ \textit{ denote the invariant measure of the kernel } P_+(x, A \times \mathbb{R}^d) \textit{ which is assumed to be irreducible and aperiodic. Let } \tau_1 = \tau_+, \tau_{j+1} = \inf\{n > \tau_j : S_{n,1} > S_{\tau_j,1}\} \text{ and }

\begin{align}
(3.1) \quad T_b &= \inf\{n \geq 1 : S_{n,1} > b\}, \\
(3.2) \quad \mu^* &= E_{\pi_+}S_{\tau_+}, \quad V_+ = \lim_{n \to \infty} n^{-1}E_{\pi_+}\{(S_{\tau_n} - n\mu^*)(S_{\tau_n} - n\mu^*)'\}.
\end{align}

Define \(\gamma_+\) and \(\hat{V}_+\) as in (2.6) but with \(\pi_+, S_{\tau_n} - S_{\tau_{n-1}}, V_+\) in place of \(\pi, \xi_n\) and \(V\).

3.1. Asymptotic distribution of \((X_{T_b}, S_{T_b})\)

We first generalize Stam's (1968) Theorem 2 for i.i.d. \(\xi_n\) to the Markov case. Let \(\sigma_1^2 = \lim_{n \to \infty} n^{-1}E_\pi(S_{\tau_n,1} - n\mu_1^*)^2\).

**Theorem 5.** Assume that \(r = 2\) in (2.3) for the ladder random walk. Then as \(b \to \infty\),

\[
\left( X_{T_b}, S_{T_b,1} - b, \sqrt{\frac{\mu_1^*}{b}} \left\{ (S_{T_b,2}, \ldots, S_{T_b,d}) - b\gamma_+ \right\} \right)
\]

converges weakly under \(P_x\) (for every \(x \in \mathcal{X}\)) to \((X, Y, W)\), where \((X, Y)\) and \(W\) are independent, \(Y\) is a positive random variable and \(X\) takes values in \(\mathcal{X}\) such that

\[
(3.3) \quad P\{X \in A, Y > y\} = \int_y^\infty P_{\pi_+}\{X_{\tau_+} \in A, S_{\tau_+,1} > u\}du/\mu_1^*
\]

for every \(A \in \mathcal{A}\) and \(y > 0\), and \(W\) is a \((d-1)\)-dimensional Gaussian vector with mean 0 and covariance matrix \(\hat{V}_+\).

For applications to nonlinear first passage problems, we replace \(S_n\) in (3.1) and in Theorem 5 by \(R_n = S_n + \Delta_n\), where \(\Delta_n\) represents some nonlinear perturbation. For the case of i.i.d. increments \(\xi_n\) and \(d = 1\), such extension has been developed by Lai and Siegmund (1977). The following theorem extends their result to the Markov case and \(d \geq 1\). Melfi (1992) has also
extended their result to the Markov case when \( d = 1 \). The proofs of Theorems 5 and 6 are given in Section 4.

**Theorem 6.** With the same assumptions as in Theorem 5, let \( \mathcal{F}_n \) be the \( \sigma \)-field generated by \( \{(X_i, Y_i), 0 \leq i \leq n\} \). Let \( \Delta_n \) be \( \mathcal{F}_n \)-measurable. Assume that for every \( x \in \mathcal{X} \),

\[
\max_{1 \leq i \leq n} |\Delta_{i,1}|/n \xrightarrow{P_x} 0, \quad \max_{1 \leq i \leq n, 2 \leq j \leq d} |\Delta_{i,j}|/\sqrt{n} \xrightarrow{P_x} 0,
\]

and that for every \( \eta > 0 \), there exist \( \delta = \delta(\eta, x) \) and \( m = m(\eta, x) \) such that

\[
P_x \{ \max_{n \leq t \leq n + \delta_n} |\Delta_{t,1} - \Delta_{n,1}| \geq \eta \} < \eta \quad \text{for all } n \geq m.
\]

Define \( \mu^*, V_+ \) by (3.2). Let \( R_n = S_n + \Delta_n \) and define \( T_b^* = \inf\{n \geq 1 : R_{n,1} \geq b\} \). Then the conclusion of Theorem 5 still holds with \( (T_b^*, R_{T_b^*}) \) in place of \( (T_b, S_{T_b}) \).

We next describe an important class of examples that motivate Theorem 6. Let \((X_n, S_n^*)\) be a Markov random walk such that \( X_n \) has stationary distribution \( \pi \) and \( \mu_\pi = E_\pi S_1^* \in \mathbb{R}^k \).

Suppose \( g : \mathbb{R}^k \rightarrow \mathbb{R} \) and \( h : \mathbb{R}^k \rightarrow \mathbb{R}^{d-1} \) are twice continuously differentiable in some neighborhood of \( \mu_\pi \). Let \( R_{n,1} = ng(S_n^*/n), (R_{n,2}, \ldots, R_{n,d})' = nh(S_n^*/n) \). Then Taylor expansions of \( g \) and \( h \) show that \( R_n \) can be expressed as \( S_n + \Delta_n \), where \( S_{n,1} = Dg(\mu_\pi)(S_n^* - n\mu_\pi) \) and \((S_{n,2}, \ldots, S_{n,d})' = Dh(\mu_\pi)(S_n^* - n\mu_\pi)\), in which \( Dg = (\partial g/\partial s_1, \ldots, \partial g/\partial s_k) \) and \( Dh = (\partial h_1/\partial s_j)_{1 \leq i \leq d-1, 1 \leq j \leq k} \). See Lai and Siegmund (1977, 1979) for certain special cases of \( g \) when \( S_n^* \) has i.i.d. increments.

### 3.2. Asymptotic expansion of the distribution of \((X_{T_b}, S_{T_b})\)

Theorem 3 (or more precisely, its variant Theorem 3* in Section 4.2) can be applied to provide a higher-order refinement of the weak convergence result in Theorem 5. In the case \( d = 2 \) and i.i.d. \( \xi_n \), such refinement has been developed in Section 4 of Keener (1988). For Markov-dependent \( \xi_n \), first consider the case where the state space is finite. In this case we can set \( \mathcal{A} = \{x\} \) in the \( U^A_\nu \) defined in (2.8) for \( x \in \mathcal{X} \). For a general state space \( \mathcal{X} \), we shall assume that there exists a \( \sigma \)-finite measure \( M \) on \((\mathcal{X}, \mathcal{A})\) such that for all \( x \in \mathcal{X} \), the probability measure \( P_x \) on \((\mathcal{X}, \mathcal{A})\) defined by \( P_x(A) = P(X_1 \in A | X_0 = x) \) is absolutely continuous with respect to \( M \) (see Section 4 of Fuh and Lai (1998)). This condition clearly implies that the measure \( P_+(x, \cdot \times \mathbb{R}^d) \) on \((\mathcal{X}, \mathcal{A})\) is also absolutely continuous with respect to \( M \) for every \( x \in \mathcal{X} \). Moreover, for fixed \( B \in \mathcal{B} \), the measure \( m_B \) on \((\mathcal{X}, \mathcal{A})\) defined by \( m_B(A) = U_{x_0}^A(B) \) is absolutely continuous with respect to \( M \), with density function \( u(x; x_0, B) \) so that \( U_{x_0}^A(B) = \int \mathbb{A} u(x; x_0, B)dM(x) \). Likewise, the renewal measure \( U_+ \) associated with the ladder random walk initialized at \((x_0, 0)\) also has density function \( u_+(\cdot; x_0, B) \) with respect to \( M \), i.e., \( U_+^A(B) = \int \mathbb{A} u_+(x; x_0, B)dM(x) \). A variant of Theorem 3, to be given in Section 4.2 as Theorem 3*, shows that \( u(x; x_0, B) \) (or \( u_+(x; x_0, B) \)) can be approximated by \( p(x)\psi_1^{x_0}(s) \),
where \( p(\cdot) \) is the density function of \( \pi \) (or \( \pi_+ \)) with respect to \( M \) and \( \psi_2^{x_0}(s) \) is given for \( s_1 > 0 \) by (2.7), with \( k = 1 \) and the coefficients of the polynomial \( \omega_1(s/\sqrt{s_1}) \) depending on \( x, x_0 \) (and with \( \gamma, \mu, \hat{V} \) replaced by \( \gamma_+, \mu^+, \hat{V}_+ \) in the case of the ladder random walk). In Section 4 we make use of this approximation and an argument similar to the proof of Theorem 4.1 in Keener (1988) to prove the following result, in which we use \( \omega_1(s/\sqrt{s_1}; x, x_0) \) (associated with the ladder random walk) to highlight the dependence of the coefficients of the polynomial on the initial state \( x_0 \) and the terminal state \( x \).

**Theorem 7.** Suppose \( r > 3 \) in (2.3) for the ladder random walk, which is also assumed to be strongly nonlattice. Let \( 0 < \delta < \min(1, r - 3), u > 0 \) and \( \alpha > 0 \). Then for every \( x_0 \in \mathcal{X}, \) as \( b \to \infty, \)

\[
P_{x_0}\{X_{T_b} \in A, S_{T_{b,1}} > b + u, \sqrt{\mu_1^+}/b((S_{T_{b,2}}, \ldots, S_{T_{b,d}}) - b\gamma_+) \in C\} =\]

\[
\frac{1}{\mu_1^+} \int_{\mathcal{X}} \left( \int_{y \in A} \int_{s \in \mathbb{R}^d, s_1 > u} \int_{t = 0}^{s_1 - u} \int_{x \in C + \sqrt{\mu_1^+}/b((s_1 - b\gamma_+) - (s_2, \ldots, s_d)) \sqrt{(2\pi)(d-1)/2} \text{det} \hat{V}_+}
\times [1 + \omega_1(x; x, x_0)/\sqrt{b}]dzdtP_+(x, dy \times ds) \right) d\pi_+(x) + o(b^{-(1+\delta)/2})
\]

uniformly in \( A \in \mathcal{A} \) and \( C \in \mathcal{B}_\alpha. \)

If one ignores terms of the order \( O(1/\sqrt{b}) \) in the integral in Theorem 7, then the integral reduces via integration by parts to

\[
P\{W \in C\} \int_{\mathcal{X}} \int_{y \in A} \int_{s \in \mathbb{R}^d, s_1 > u} (s_1 - u)P_+(x, dy \times ds) d\pi_+(x) = P\{W \in C\} \int_u^\infty P_{\pi_+}(X_{\tau_+} \in A, S_{\tau_+,1} > v) dv,
\]

where \( W \) is a \((d - 1)\)-dimensional Gaussian vector with mean 0 and covariance matrix \( \hat{V}_+ \), as in Theorem 5.

### 3.3. Asymptotic expansion of the variance of \( T_b \)

Since \( T_b \) only involves \( S_{n,1} \), we can assume without loss of generality that \( d = 1 \) to analyze \( \text{Var}T_b \) so that we can write \( S_{n} \) instead of \( S_{n,1} \) to simplify the notation in the following theorem. When the \( \xi_n \) are i.i.d., Smith (1959) derived an asymptotic expansion of \( \text{Var}T_b \) under the assumption that the \( \xi_n \) are nonnegative, which Lai and Siegmund (1979) subsequently removed. The following theorem extends the result to Markov random walks. As in Section 3.2, we shall assume that there exists a \( \sigma \)-finite measure \( M \) on \((\mathcal{X}, \mathcal{A})\) such that \( P_x(X_1 = \cdot) \) is absolutely continuous with respect to \( M \) for all \( x \in \mathcal{X} \). Define \( \tau_- = \inf\{n \geq 1 : S_n \leq 0\} \) and \( \tau_+ \) as before.

Then for every \( x \in \mathcal{X} \), there exists \( p_-(x; \cdot) \) such that \( P(X_{\tau_-} \in A|X_0 = x) = \int_A p_-(x; y) dM(y) \) for all \( A \in \mathcal{A} \). Suppose \( \{X_n, n \geq 0\} \) has a stationary distribution \( \pi \) which has a positive density function \( p \) with respect to \( M \). As in Section 4 of Fuh and Lai (1998), consider the time-reversal
(dual) process \( \{(\bar{X}_n, \bar{S}_n), n \geq 0\} \) and define \( \bar{p}_- \) and \( \bar{G}_- \) for the dual process in the same way as \( p_- \) and \( G_- \) are defined for \( \{(X_n, S_n), n \geq 0\} \). Let \( \bar{G}_-(x, y; B) = G_-(y, x; B) p(y)/p(x), \) \( \bar{U}_- = \sum_{n=0}^\infty \bar{G}_-^{* n} \), where \( * \) denotes convolution of two transition kernels \( F_1(x, y; \cdot) \) and \( F_2(x, y; \cdot) \) as defined in (4.3) of Fuh and Lai (1998), with \( \bar{G}_-^{* 1} = \bar{G} \) and \( \bar{G}_-^{* 0} \) being the kernel that puts all its mass at 0.

**Theorem 8.** Suppose that (2.1) and (2.2) hold for both \( \{(X_n, S_n), n \geq 0\} \) and its ladder random walk, and that (2.3) holds with \( r = 2 \) for \( \{(X_n, S_n), n \geq 0\} \) and with \( r > 3 \) for the ladder random walk, which is also assumed to be strongly nonlattice. Let \( \mu = E_x S_1, \) \( \sigma^2 = \lim_{n \to \infty} n^{-1} E_x (S_n - n \mu)^2 \) and \( \mu^* = E_{x^+} S_{r^+} \). Then as \( b \to \infty, \)

\[
E_x(T_b) = \mu^{-1} \left( b + \frac{1}{2} E_{x^+} S_{r^+}^2 / \mu^* - \int_{\mathcal{X}} g_1(y) d\pi_+(y) + g_1(x) \right) + o(1),
\]

\[
Var_x(T_b) = \mu^{-3} \sigma^2 b + \mu^{-2} K + o(1)
\]

for every \( x \in \mathcal{X} \), where \( g_1: \mathcal{X} \to \mathbb{R} \) solves the Poisson equation

\[
E_x g_1(X_1) - g_1(x) = E_x \xi_1 - E_x \xi_1
\]

for almost every (with respect to \( M \)) \( x \in \xi \), with \( E_x g_1(X_1) = 0 \), \( g_2 = h - g_1^2 \) and \( h \) is a solution of the Poisson equation

\[
E_x h(X_1) - h(x) = E_x \{\xi_1 - \mu - g_1(X_1) + g_1(x)\}^2 - E_x \{\xi_1 - \mu - g_1(X_1) + g_1(X_0)\}^2
\]

for almost every \( x \), with \( E_x h(X_1) = 0 \), and

\[
K = \frac{\sigma^2}{\mu} \left\{ \frac{E_{x^+} S_{r^+}^2}{2 \mu^*} - \int_{\mathcal{X}} g_1(y) d\pi_+(y) + g_1(x) \right\} - \frac{E_{x^+} S_{r^+}^3}{3 \mu^*} + \left( \frac{E_{x^+} S_{r^+}^2}{2 \mu^*} \right)^2
\]

\[
- \left\{ \int_{\mathcal{X}} g_1(y) d\pi_+(y) - g_1(x) \right\}^2 + \int_{\mathcal{X}} g_2(y) d\pi_+(y) - g_2(x)
\]

\[
- 2 \int_0^\infty \int_{\mathcal{X}} s g_1(z) P_{r^+}(X_{r^+} \in dz, S_{r^+} > s) ds / \mu^*
\]

\[
+ \frac{2 \mu}{\mu^*} \int_0^\infty \left\{ E_{x^+} (S_{T_v} - v) - \frac{E_{x^+} S_{r^+}^2}{2 \mu^*} \right\} \int_{\mathcal{X}} \int_{\mathcal{X}} \bar{U}_-(z, y; (-v, 0)) dM(y) d\pi_+(z) dv.
\]

Conditions (2.1) and (2.2) for \( \{(X_n, S_n), n \geq 0\} \) ensure not only the existence of solutions to the Poisson equations (3.8) and (3.9) but also

\[
\sup_x |g_j(x)|/w(x) < \infty \text{ for } j = 1, 2;
\]

see Theorem 17.4.2 and Lemma 15.2.9 of Meyn and Tweedie (1993). Moreover, under (2.3) with \( r = 2 \), \( \mu \) and \( \sigma^2 \) are well defined. Conditions (2.1)-(2.3) (with \( r > 3 \)) for the associated ladder random walk enable us to apply Theorem 4 to the ladder random walk that is associated with \( \mu^* \) and terms involving \( \pi_+, X_{r^+} \) and \( S_{r^+} \) in the asymptotic formulas of Theorem 8. Note
that in Examples 1-4 both \(\{(X_n, S_n), n \geq 0\}\) and its ladder random walk satisfy (2.1)-(2.3) with \(r > 3\). The proof of Theorem 8 uses Wald’s equations for \(E_x S_{T_b} \) and \(E_x(S_{T_b} - \mu T_b)^2\), which involve the functions \(g_1\) and \(g_2\), together with the following lemma, which uses the same notation and assumptions as in the paragraph preceding Theorem 8 and which is an extension of Theorem 4 of Lai and Siegmund (1979) from i.i.d. to Markov dependent \(\xi_n\).

**Lemma 1.** Suppose the Markov random walk \(\{(X_n, S_n), n \geq 0\}\) (with \(d = 1\)) is nonlattice and \(E_x|S_1| < \infty, E_x S_1 > 0\). Then for any \(u > 0, x \in \mathcal{X}\) and \(A \in \mathcal{A}\),

\[
\lim_{b \to \infty} \sum_{n=0}^{\infty} P_x \{T_b > n, X_n \in A, S_n > b - u\} = (E_{\pi^+ S_{\tau_+}})^{-1} \int_{u}^{0} \int_{A} \left\{ \int_{A} \tilde{U}_{-}(z, y; (s, 0))dM(y) \right\} d\pi^+ (z)ds.
\]

*Proof.* First note that for any \(z \in \mathcal{X}\),

\[
P_x \{T_b > n, X_n \in A, S_n > b - u\} = \sum_{k=0}^{n} \int_{(b-u, b)} P_x \{S_i < S_k \text{ for all } i < k, S_k \in dt, S_j \leq S_k \text{ for all } k < j \leq n, S_n - S_k > b - u - t, X_n \in A\} = \sum_{k=0}^{n} \int_{(0,u)} \int_{\mathcal{X}} P_x \{S_i < S_k \text{ for all } i < k, S_k \in b - ds, X_k \in dx\} \times P_x \{\tau_+ > n - k, S_{n-k} > s - u, X_{n-k} \in A\}.
\]

Let \(\tau_m\) be the \(m\)th ascending ladder epoch. Then \(\sum_{k=0}^{\infty} P_x \{S_i < S_k \text{ for all } i < k, S_k \in b - ds, X_k \in dx\} = \sum_{m=0}^{\infty} P_x \{S_{\tau_m} \in b - ds, X_{\tau_m} \in dz\}\). As shown by Fuh and Lai (1998, page 576), \(\sum_{j=0}^{\infty} P_x \{\tau_+ > j, S_j \in B, X_j \in A\} = \int_{A} \tilde{U}_{-}(z, y; B)dM(y)\). Combining these representations with the Markov renewal theorem (cf. Alsmeyer (1994)) yields

\[
\sum_{n=0}^{\infty} P_x \{T_b > n, X_n \in A, S_n > b - u\} = \sum_{j=0}^{\infty} \int_{(0,u)} \int_{\mathcal{X}} \sum_{m=0}^{\infty} P_x \{S_{\tau_m} \in b - ds, X_{\tau_m} \in dz\} P_x \{\tau_+ > j, S_j > s - u, X_j \in A\} = \int_{(0,u)} \int_{\mathcal{X}} \left\{ \int_{A} \tilde{U}_{-}(z, y; (s - u, 0))dM(y) \right\} \sum_{m=0}^{\infty} P_x \{S_{\tau_m} \in b - ds, X_{\tau_m} \in dx\} \rightarrow (E_{\pi^+ S_{\tau_+}})^{-1} \int_{u}^{0} \int_{\mathcal{X}} \left\{ \int_{A} \tilde{U}_{-}(z, y; (s - u, 0))dM(y) \right\} d\pi^+ (z)ds \text{ as } b \to \infty.
\]

*Proof of Theorem 8.* By Corollary 1 and Theorem 4 of Fuh and Zhang (2000), we have Wald’s equations

\[
E_x S_{T_b} = \mu E_x T_b + E_x g_1(X_{T_b}) - g_1(x),
\]

\[
E_x(S_{T_b} - \mu T_b)^2 = \sigma^2 E_x T_b - 2E_x \{(S_{T_b} - \mu T_b)g_1(X_{T_b})\} + E_x g_2(X_{T_b}) - g_2(x).
\]

12
Moreover, $g_1$ and $g_2$ satisfy (3.10); see the paragraph proceeding Lemma 1. In view of (3.10) and the finiteness of $\int w(x) d\pi(x)$, we can apply Corollary 2.2 of Alsmeyer (1994) to show that

$$\lim_{b \to \infty} E_x g_j(X_{T_b}) = \int g_j(y) d\pi_+(y) \text{ for } j = 1, 2.$$  

(Actually (1.5) of Alsmeyer (1994) is not properly stated and should be restated as that $E_x g(X_{T_b}, S_{T_b} - b)$ converges to $E_g(X, Y)$ as $b \to \infty$, where $(X, Y)$ has the limiting distribution (3.3) of $(X_{T_b}, S_{T_b} - b)$.) We next show that for $i = 1, 2$,

$$\lim_{b \to \infty} E_x (S_{T_b} - b)^i = (E_x S_{T_b})^{-1} \int_0^\infty s^i P_{\pi_+}(S_{T_b} > s) ds.$$  

To prove (3.14), we can assume without loss of generality that $\xi_n \geq 0$ a.s. by working with the ladder random walk. We can then proceed as in Siegmund (1985, pp. 167, 170 and 171). In particular, analogous to his representation (8.11) for i.i.d. $\xi_n$, we now have

$$P_x\{S_{T_b} - b > s\} = \int_\mathcal{X} \int_{[0,b]} P_y\{\xi_1 > b + s - t\} u(y; x, dt) dM(y),$$  

where $u(y; x, B)$ is defined in the paragraph preceding Theorem 7.

Since $E_x S_{T_b} = b + E_x (S_{T_b} - b)$, it follows from (3.11), (3.13) and (3.14) that (3.6) holds. Moreover, by (3.12),

$$\mu^2 \text{Var}_x T_b = E_x (\mu T_b - S_{T_b} + T_b - \mu E_x T_b)^2$$  

$$= E_x (S_{T_b} - \mu T_b)^2 + E_x (S_{T_b} - \mu E_x T_b)^2 - 2E_x [(S_{T_b} - \mu T_b)(S_{T_b} - \mu E_x T_b)]$$  

$$= \sigma^2 E_x T_b - 2E_x \{S_{T_b} - \mu T_b\} g_1(X_{T_b}) + E_x g_2(X_{T_b}) - g_2(x)$$  

$$+ E_x (S_{T_b} - b + b - \mu E_x T_b)^2 - 2E_x [(S_{T_b} - b + b - \mu T_b)(S_{T_b} - b + b - \mu E_x T_b)].$$

Analogous to (3.15), we also have

$$P_x\{S_{T_b} - b > s, X_{T_b} \in dz\} = \int_\mathcal{X} \int_{[0,b]} P_y\{\xi_1 > b + s - t, X_1 \in dz\} u(y; x, dt) dM(y),$$  

and a straightforward modification used to prove (3.14) can be used to show that as $b \to \infty$,

$$E_x \{S_{T_b} - b\} g_1(X_{T_b}) \longrightarrow \int_0^\infty \int_{\mathcal{X}} s g_1(z) P_{\pi_+} \{X_{T_b} - s \in dz, S_{T_b} > s\} ds/\mu^*.$$  

Putting (3.6), (3.8), (3.14) and (3.17) into (3.16) yields

$$\mu^2 \text{Var}_x T_b = \sigma^2 b/\mu + \sigma^2 E_{\pi_+} S_{T_b} \mu^2 - \sigma^2 \{\int g_1(y) d\pi_+(y) - g_1(x)\}/\mu$$  

$$+ \int g_2(y) d\pi_+(y) - g_2(x) + (E_{\pi_+} S_{T_b}^2/2\mu^*)^2 - \{\int g_1(y) d\pi_+(y) - g_1(x)\}^2$$  

$$- 2 \int_0^\infty \int_{\mathcal{X}} s g_1(z) P_{\pi_+} \{X_{T_b} \in dz, S_{T_b} > s\} ds/\mu^* - E_{\pi_+} S_{T_b}^3/3\mu^*$$  

$$+ 2\mu E_{\pi_+} [(T_b - E_x T_b)(S_{T_b} - b)] + o(1).$$
Let $\psi(v) = E_{\pi_+} (S_{T_b} - v) - E_{\pi_+} S^2_{T_b} / 2 \mu^*$. It remains to show that as $b \to \infty$,

$$
(3.18) \quad \frac{E}{X} \left\{ (T_b - E_z T_b) (S_{T_b} - b) \right\} \rightarrow \int_0^\infty \psi(v) \int_X \tilde{U}(z, y; (-v, 0)) dM(y) d\pi_+(z) dv / \mu^*.
$$

As in Eq. 5.7 of Lai and Siegmund (1979) for the case of i.i.d. $\xi_n$, the left hand side of (3.18) can be expressed as

$$
(3.19) \quad \sum_{n=0}^\infty \int_0^\infty \int_X \left\{ E_z (S_{T_n} - v) - E_z (S_{T_n} - b) \right\} P_z \{ T_n > n, X_n \in dz, S_n \in b - dv \}.
$$

Making use of Theorem 4 applied to the ladder random walk, it will be shown that as $v \to \infty$,

$$
(3.20) \quad E_z (S_{T_v} - v) = E_{\pi_+} S^2_{T_b} / 2 \mu^* + o(v^{-(r-2)})
$$

uniformly in $z \in X$. Since $r > 2$, (3.18) follows from (3.19), (3.20) and Lemma 1. To prove (3.20), we again assume that $\xi_n \geq 0$ a.s. so that $\pi_+ = \pi$ and $\mu^* = \mu$. From (3.15) it follows that

$$
(3.21) \quad E_z (S_{T_b} - b) - (E_{\pi_+} (\xi_1^2) / 2 \mu = \int_X \int_{(0,b]} \{ \int_0^\infty P_z (\xi_1 > s) ds \} u(y, x, b - dt) - p(y) dt / \mu dM(y),
$$

where $p(\cdot)$ is the density function of $\pi$ with respect to $M$. Since $\sup_y E_y \xi_1^2 / w(y) < \infty$ by (2.3),

$$
(3.22) \quad \sup_y \int_0^\infty P_y (\xi_1 > s) ds / w(y) = O(t^{-(r-1)}).
$$

The desired conclusion then follows by decomposing the interval $(0, b]$ in (3.21) as the union $(\cup_{1 \leq j \leq [b/2]} (j - 1, j)) \cup ([b/2], b]$ and applying Theorem 4 to each of the subintervals $(j - 1, j]$, $1 \leq j \leq j_b := [b/2]$, and (3.22) to the subinterval $(j_b, b]$, noting that $\int w(y) d\pi(y) < \infty$. 

4. Proof of Theorems 1-7

To prove Theorems 1-4, we evaluate the Fourier transform of the renewal measure $U^A$. As in Carlsson (1982, 1983) and Carlsson and Wainger (1982), we perform Fourier inversion of the Fourier transform as a generalized function. We refer the reader to Gelfand and Shilov (1964), Schwartz (1966) and Strichartz (1994) for the basic theory; in particular, the following notation and concepts will be used.

A test function $\varphi(s) = \varphi(s_1, \cdots, s_d)$ is an infinitely differentiable function that vanishes outside a bounded region in the $d$-dimensional space $\mathbb{R}^d$. Let $\mathcal{D}$ denote the linear space of all test functions, and $\mathcal{D}'$ the space of linear functionals on $\mathcal{D}$. A sequence $\varphi_n \in \mathcal{D}$ is said to converge to zero if $\varphi_n$ and all its derivatives converge to 0 uniformly and vanish outside.
a common bounded subset of $\mathbb{R}^d$. A \textit{generalized function} is a continuous linear functional on $\mathcal{D}$. A function $f$ defined on $\mathbb{R}^d$ for which $\int f(s) \varphi(s) ds$ is absolutely convergent for any $\varphi \in \mathcal{D}$ is called \textit{locally integrable}. A $C^\infty$ function $f$ on $\mathbb{R}^d$ is of class $\mathcal{T}$ if $f$ and all its partial derivatives are \textit{rapidly decreasing} in the sense that they are of order $O(|s|^{-a})$ as $|s| \to \infty$, for every $a > 0$. Linear functionals on $\mathcal{T}$ are called \textit{tempered distributions}, and $\mathcal{T}'$ denotes the set of all tempered distributions. The Fourier transform $\hat{\varphi}$ of a function $\varphi \in \mathcal{D}$ is defined by $\hat{\varphi}(\theta) = \int \varphi(s) \exp(i\theta^t s) ds$ for $\theta \in \mathbb{R}^d$. The Fourier transform of a generalized function $f$ is the linear functional $\hat{f}$ defined on the space $\{\psi : \psi$ is the Fourier transform of some $\varphi \in \mathcal{D}\}$ by $(2\pi)^d(f, \varphi) = (\hat{f}, \hat{\varphi})$ for all $\varphi \in \mathcal{D}$.

Throughout the sequel, let $\tilde{S}_n = (S_{n,1}, \tilde{S}_{n,2}, \ldots, \tilde{S}_{n,d})'$ when $d > 1$. In the case of i.i.d. $\xi_n$, $\sum_{n=0}^\infty E(e^{i\theta^t \tilde{S}_n}) = \sum_{n=0}^\infty \lambda^n(\theta)^{-1} = (1 - \lambda(\theta))^{-1}$, where $\lambda(\theta) = Ee^{i\theta^t \xi}$, and Theorems 1-4 have been derived for this special case via the generalized function $(1 - \lambda(\theta))^{-1}$. For Markov dependent $\xi_n$, the Fourier transform of the renewal measure $U^A$ does not have such explicit representation. Nevertheless the following two key lemmas in the Markov case can be used in lieu of the explicit representation for the analysis of the Fourier transform

$$
(4.1) \quad f_A(\theta) = \sum_{n=0}^\infty E_x(e^{i\theta^t \tilde{S}_n} 1_{\{X_n \in A\}}).
$$

Lemma 2 gives a decomposition of $f_A(\theta)$ in terms of $e(\theta)(1 - \lambda(\theta))^{-1} + \eta(\theta)$ for $|\theta| \leq \delta$, with $\delta > 0$ sufficiently small, where $\lambda(0) = 1$ and $\lambda(\theta), e(\theta), \eta(\theta)$ are smooth functions for $|\theta| \leq \delta$.

Lemma 3 shows that $f_A(\theta)$ and its partial derivatives up to order $[r]$ are bounded in the region $\{\theta : |\theta| \geq \delta\}$.

The decomposition in Lemma 2 follows from the spectral decomposition of certain linear operators on $B$, using the same notation and assumptions as in the first paragraph of Section 2. For $d \times 1$ vectors $\theta$, define the linear operators $P_{\theta}$, $P$, $\nu_\theta$ and $Q$ on $B$ by

$$
(4.2) \quad (P_{\theta} h)(x) = \int h(y) e^{i\theta^t y} P(x, dy \times ds) = E_x\{h(X_1) e^{i\theta^t S_1}\},
$$

$$
(4.3) \quad (P h)(x) = \int h(y) P(x, dy \times ds) = E_x h(X_1),
$$

$$
(4.4) \quad \nu_\theta h = E_x\{h(X_0) e^{i\theta^t S_0}\}, \quad Q h = \int h(y) d\mu(y).
$$

For the case of $w \equiv 1$ and $\xi_n = g(X_n)$, Nagaev (1957) and Jensen (1987) have shown that there exists sufficiently small $\delta > 0$ such that for $|\theta| \leq \delta$, $B = B_1(\theta) \oplus B_2(\theta)$ and

$$
(4.5) \quad P_{\theta} Q_{\theta} h = \lambda(\theta) Q_{\theta} h \quad \text{for all } h \in B,
$$

where $B_1(\theta)$ is a one-dimensional subspace of $B$, $\lambda(\theta)$ is the eigenvalue of $P_{\theta}$ with corresponding eigenspace $B_1(\theta)$, $Q_{\theta}$ is the parallel projection of $B$ onto the subspace $B_1(\theta)$ in the direction of $B_2(\theta)$. Extension of their argument to general $w$ and $\xi_n$ satisfying the assumptions of Section
Lemma 2. Under (2.1)-(2.3) and (2.9), let \( h \in \mathcal{B} \) and \( |\theta| \leq \delta \),

(i) \( E\nu\{e^{i\theta S_n}h(X_n)\} = \lambda^n(\theta)\nu_\theta Q_\theta h + \nu_\theta \mathbf{P}_\theta^n (I - Q_\theta)h \). Moreover, there exist \( 0 < \delta^* \leq \delta \), \( 0 < \gamma < 1 \) and \( K > 0 \) such that for \( |\theta| \leq \delta^* \), \( \lambda(\theta), \nu_\theta Q_\theta h \) and \( \sum_{n=0}^{\infty} \nu_\theta \mathbf{P}_\theta^n (I - Q_\theta)h \) have continuous partial derivatives of order \( [r] \), and \( |\nu_\theta \mathbf{P}_\theta^n (I - Q_\theta)h| \leq K\|h\|_w|\theta|\gamma^n \) for all \( n \geq 1 \).

Furthermore,

\[ \lambda(0) = 1, \ \nabla \lambda(0) = i\Gamma u, \ \nabla^2 \lambda(0) = -\Gamma \nabla \Gamma^r. \]

(ii) Define \( f_A(\theta) \) by (4.1) and let \( h_A(\theta) = 1_{\{x \in A\}} \). Then for \( 0 < |\theta| \leq \delta^* \),

\[ f_A(\theta) = (1 - \lambda(\theta))^{-1}\nu_\theta Q_\theta h_A + \eta(\theta), \]

where \( \eta(\theta) \) has continuous partial derivatives of order \( [r] \) and \( \eta(\theta) = O(|\theta|) \) as \( \theta \to 0 \).

Lemma 3. Suppose \( \{(X_n, S_n), n \geq 0\} \) is a strongly nonlattice Markov random walk satisfying (2.1)-(2.3) and (2.9) for some \( r \geq 3 \). Then for every \( A \in \mathcal{A} \), \( a > 0 \) and \( (j_1, \ldots, j_d) \) with \( \min_{k \leq d} j_k \geq 0 \) and \( \sum_{k=1}^d j_k \leq r \),

\[ \sup_{|\theta| > a} \frac{1}{\sum_{n=0}^{\infty}} |E\nu\left( S_{n,1}^{j_1} \cdots S_{n,d}^{j_d} e^{i\theta S_n} 1_{\{X_n \in A\}} \right)| < \infty. \]

Consequently, \( f_A(\theta) \) and its partial derivatives of order not exceeding \( r \) are bounded for \( |\theta| \geq a \), where \( f_A \) is defined by (4.1).

Proof. We first consider the case \( j_1 = \cdots = j_d = 0 \). Let \( g(x, y) = E(e^{i\theta \xi_1}|X_0 = x, X_1 = y) \). For \( n \geq 8 \), let \( l = \left\lceil (\log n)^2 \right\rceil \), \( k_n = \left\lceil n/l \right\rceil - 1 \) and \( J_n = \{ j l + 1 \leq j \leq k_n \} \).

Then

\[ |E_x(e^{i\theta S_n})| = |E_x\left[ E\left( \sum_{t=1}^n e^{i\theta \xi_t} \left\{ (X_t, X_{t+1}) : t \in J_n \right\} \cup \{ \xi_t : 0 \leq t \leq n, t \notin J_n \} \right) \right]| \]

\[ \leq |E_x\left\{ \prod_{j \in J_n} E(e^{i\theta \xi_j}|X_t, X_{t+1}) \right\}| = |E_x\left\{ \prod_{j=1}^{k_n} g(X_{j_1}, X_{j_1+1}) \right\}|, \]

noting that the \( \xi_t \) are conditionally independent given \( (X_t, X_{t+1}), t \in J_n \). By Theorem 16.1.5 of Meyn and Tweedie (1993), there exist \( A > 0 \) and \( 0 < \gamma < 1 \) such that

\[ |E_x\left\{ \prod_{j=1}^{k_n} g(X_{j_1}, X_{j_1+1}) \right\} - E_x\left\{ \prod_{j=1}^{k_n-1} g(X_{j_1}, X_{j_1+1}) \} E_x g(X_{k_n}, X_{k_n+1}) \right\}| \leq A\gamma^{l-1} w(x), \]

and therefore by induction

\[ |E_x\left\{ \prod_{j=1}^{k_n} g(X_{j_1}, X_{j_1+1}) \right\} - \prod_{j=1}^{k_n-1} E_x h(X_{j_1}) | \leq (k_n - 1) A\gamma^{l-1} w(x), \]
where \( h(z) = E_z(g(z, X_1)) \) so that \( E_x g(X_{j_1}, X_{j_1+1}) = E_x h(X_{j_1}) \). Note that \( \int h(z) d\pi(z) = \int E_x(e^{i\theta \xi_1}) d\pi(z) := \varphi(\theta) \). By (2.4), \( |E_x h(X_{j_1}) - \int h(z) d\pi(z)| \leq C \rho^{j_1} w(x) \) with \( 0 < \rho < 1 \), and therefore

\[
(4.9) \quad \prod_{j=1}^{k_n} |E_x h(X_{j_1})| \leq \prod_{k_n/2 \leq j \leq k_n} |E_x h(X_{j_1})| \leq \prod_{k_n/2 \leq j \leq k_n} \{ |\varphi(\theta)| + C \rho^{j_1} w(x) \}.
\]

By the strong nonlattice property (2.5), there exists \( 0 < \zeta < 1 \) such that \( \sup_{|\theta| \geq a} |\varphi(\theta)| \leq \zeta \). For \( k_n/2 \leq j \leq k_n \) and all sufficiently large \( n \), \( j_1 \geq \frac{1}{2} l([n/l] - 1) \geq n/3 \) and therefore \( \rho^{j_1} \leq \rho^{n/3} \) (since \( 0 < \rho < 1 \)). By Markov's inequality,

\[
(4.10) \quad \nu \{ x : C \rho^{n/3} w(x) \geq (1 - \zeta)/2 \} \leq 2(1 - \zeta)^{-1} C \rho^{n/3} \int w(x) d\nu(x).
\]

From (4.7)-(4.10), it follows that for \( |\theta| \geq a \) and all sufficiently large \( n \),

\[
\begin{align*}
&\int |E_x(e^{i\theta S_n})| d\nu(x) \leq (k_n - 1) A \gamma^l - 1 \int w(x) d\nu(x) \\
&+ \int C \rho^{n/3} w(x) \zeta \leq 2(1 - \zeta)^{-1} C \rho^{n/3} \int w(x) d\nu(x) \\
&= O(k_n \gamma^l) + O((1 + \zeta)^{2} k_n^{2}) + O(\rho^{n/3}).
\end{align*}
\]

Since \( l \sim (\log n)^2 \) and \( k_n \sim n (\log n)^{-2} \) while \( \gamma, \rho \) and \( \zeta \) are positive numbers \( < 1 \), we obtain (4.6) in the case \( j_1 = \ldots = j_d = 0 \).

To fix the ideas and simplify the notation when \( \min(j_1, \ldots, j_d) > 0 \), we focus on the special case \( d = 2 \) and \( j_1 = 1, j_2 = 1 \), for which

\[
S_{n,1}^2 = \sum_{t=1}^{n} \sum_{k=1}^{n} \xi_t \xi_{k,2} + 2 \sum_{1 \leq t < j \leq n} \sum_{k=1}^{n} \xi_t \xi_{j,1} \xi_{k,2}.
\]

There are \( O(n^2) \) summands on the right-hand side and we need only prove that the product of each summand and \( e^{i\theta S_n} \) has expected value of order \( O(n \gamma^l) \), where \( l \sim (\log n)^2 \) as before. The proof proceeds by using an argument similar to that above together with extensions of Lemma 15.2.9 and Theorem 16.1.5 of Meyn and Tweedie (1993) given in Lemma 4 below. In particular, consider \( \xi_t \xi_{j,1} \xi_{k,2} \) with \( k > j \) and \( j - t > [n/l] + 2l \). Let \( I_n = \{ t + ml : 1 \leq m \leq [n/l] \} \). Then analogous to (4.7), \( |E_x(\xi_t \xi_{j,1} \xi_{k,2} e^{i\theta S_n})| \) is majorized by

\[
(4.11) \quad |E_x \left( E(\xi_t \xi_{j,1} \xi_{k,2} | X_{t-1}) \prod_{\tau \in I_n} E(e^{i\theta \xi_{\tau}} | X_{\tau}, X_{\tau+1}) E(\xi_{j,1} \xi_{k,2} e^{i\theta (\xi_{\tau} + \xi_{k})} | X_{\tau-1}) \right) |.
\]

Note that by (2.3) with \( r > 3 \), \( E_y|\xi_t| \leq (E_y|\xi_t|^r)^{1/r} \leq A w^{1/3}(y) \leq A w^{1/3} \) for sufficiently large \( A \). Making use of Lemma 4(i) below, it can be shown that \( E(|\xi_{j,1} \xi_{k,2}||X_{j-1}) \leq A w^{1/3}(X_{j-1}) \), with \( A \) sufficiently large. Hence by Lemma 4(ii) and an induction argument as in (4.8), we obtain that (4.11) is majorized by

\[
|E_x(\xi_t \xi_{j,1} \xi_{k,2} e^{i\theta \xi_{\tau}}) \prod_{\tau \in I_n} E_x(e^{i\theta \xi_{\tau}}) E_x(\xi_{j,1} \xi_{k,2} e^{i\theta (\xi_{\tau} + \xi_{k})}) + O([n/l] \gamma^l w(x)).
\]
Since $E_x[\xi_j,1,\xi_{j,k}] \leq \tilde{A}w^{2/3}(x)$ and $E_x[\xi_{l,1}] \leq \tilde{A}w^{1/3}(x)$, we can then use the same argument as in the preceding paragraph to bound $\int_{\omega^{n/3}w(x) < (1-\epsilon)/2} |E_x(\xi_{l,1,1,1,1,2}e^{i\theta S_n})|\nu(x)$. Moreover, by Lemma 4(i), $E_x[\xi_{l,1,1,1,1,1,2}] \leq Aw^{3/r}(x)$, and therefore

$$
\int_{Cw(x) \geq \rho^{-n/3}(1-\epsilon)/2} E_x[\xi_{l,1,1,1,1,1,2}]\nu(x)
\leq \left(2\rho^{n/3}/(C(1-\epsilon))\right)^{(r-3)/r} \int_{Cw(x) \geq \rho^{-n/3}(1-\epsilon)/2} Aw(x)\nu(x) = O(\rho^{r-3/3r}).
$$

In general, given $k$ indices $t_1, \ldots, t_k$ between 1 (= $t_0$) and $n$ (= $t_{k+1}$), there exists $j$ such that $t_j - t_{j-1} \geq n/(k+1)$. For large $n, n/(k+1) > [n/l] + 2l$ with $l = [(\log n)^2]$. We can therefore proceed as before, taking $\{t_j, t_{j+1}, \ldots, t_k\}$ as one group, $\{t_1, \ldots, t_{j-1}\}$ as another group and $I_n = \{t_{j-1} + ml : 1 \leq m \leq [n/l]\}$. □

The following lemma, which generalizes corresponding results in Meyn and Tweedie (1993) from $\alpha = 1/2$ to $0 \leq \alpha \leq 1$, has been used in the preceding proof.

Lemma 4. Let $\{X_n, n \geq 0\}$ be an irreducible, aperiodic Markov chain satisfying (2.1). Let $0 \leq \alpha \leq 1$.

(i) There exist $0 < \lambda_\alpha < 1$ and $L_\alpha > 0$ such that $E_xw^\alpha(X_1) \leq \lambda_\alpha w^\alpha(x) + L_\alpha$ for all $x \in \mathcal{X}$. Consequently, for all $x \in \mathcal{X}$ and $n \geq 1$, $E_xw^\alpha(X_n) \leq \lambda_\alpha w^\alpha(x) + L_\alpha/(1 - \lambda_\alpha)$.

(ii) There exist $0 < \rho_\alpha < 1$ and $A_\alpha > 0$ such that for any $|g| \leq w^\alpha$ and $|h| \leq w^{1-\alpha}$, $k \geq 1$, $n \geq 1$ and $x \in \mathcal{X}$, $|E_x\{g(X_k)h(X_{n+k})\} - \{E_xg(X_k)\}E_xh(X_{n+k})| \leq A_\alpha\rho_\alpha^k(1 + \rho^k w(x))$.

Proof. By Jensen's inequality and Lemma 15.2.8 of Meyn and Tweedie (1993), $E_xw^\alpha(X_1) \leq \{E_xw(X_1)\}^\alpha \leq \{\rho_\alpha(x) + L\}^\alpha$ with $0 < \rho < 1$ and $L > 0$. Moreover, $(1 + t)^\alpha \leq 1 + \alpha t$ for $t \geq 0$. Hence (i) follows. The proof of (ii) makes use of (i) and an argument similar to the proof of Theorem 16.1.5 of Meyn and Tweedie (1993). □

4.1. Proof of Theorems 1-3

Define $f_A$ by (4.1). Since $E_x\tilde{S}_{n,k} = 0$ for $2 \leq k \leq d$ and $\lambda(\theta)$ has continuous partial derivatives of order $[r]$ for $|\theta| \leq \delta^*$ (see Lemma 2), $\lambda(\theta)$ has a Taylor expansion of the form $\lambda(\theta) = 1 + i\theta_1 \mu_1 - \frac{1}{2} \theta_1^T \tilde{\theta}_{(-1)} + \sum_{j=3}^M p_j(\theta) + o_P(M)$ with $M \leq [r]$, where $\theta_{(-1)} = (\theta_2, \ldots, \theta_d)'$ and $p_j(\theta)$ is a polynomial homogenous in the sense that $p_j(\alpha^2 \theta_1, \alpha \theta_{(-1)}) = p_j(\theta)$. Since $p_j(\theta)$ is a higher-order polynomial than $i\theta_1 \mu_1 - \frac{1}{2} \theta_1^T \tilde{\theta}_{(-1)}$, the following expansion holds for $0 < |\theta| \leq \delta^*$:

$$
(1 - \lambda(\theta))^{-1} = \sum_{j=0}^{M-2} \left(- \sum_{m=3}^{M} p_m(\theta)\right)^j/(-i\theta_1 \mu_1 + \frac{1}{2} \theta_{(-1)}^T \tilde{\theta}_{(-1)})^{j-1} + \Delta_M(\theta).
$$

Proof of Theorem 1. Let $N = 1 + [(k+d-1)/2], M = 2N - d + 1$ and $\epsilon = s_1^{-L}$, with $L$ large enough. As in Carlsson and Wainger (1982), abbreviated by CW hereafter, we shall assume that the $d$-dimensional rectangle $B$ is the unit cube and that $\gamma = 0$ so that $\tilde{s} = (s_2, \ldots, s_d)'$
and \( \theta(-1) = \tilde{\theta} \). Let \( \phi_\varepsilon(x) = \varepsilon^{-d} \phi(x/\varepsilon) \), where \( \phi \) is an infinitely differentiable function with support in \( B \), and let \( B_a = [-a,a]^d \). Moreover, to simplify the notation, we shall assume as in CW that \( \mu_1 = 1 \) and \( V = I \). By Lemma 2(ii) and Fourier inversion of the Fourier transform as a generalized function (noting that \( (1 - \lambda(\theta))^{-1} \) is undefined at \( \theta = 0 \)), the convolution 

\[
(2\pi)^d (\phi_\varepsilon * 1_{B_a} * U^A_v(s)) \text{ can be expressed as}
\]

\[
(4.13) \quad \int e^{-i\theta^* \phi}(\varepsilon \theta) \hat{1}_{B_a}(\theta) f_A(\theta) d\theta = \int_{0<|\theta|\leq \delta^*} e^{-i\theta^* \phi}(\varepsilon \theta) \hat{1}_{B_a}(\theta)(1 - \lambda(\theta))^{-1} \nu_{\theta} Q_{\theta^*} h_A d\theta \\
+ \int_{0<|\theta|\leq \delta^*} e^{-i\theta^* \phi}(\varepsilon \theta) \hat{1}_{B_a}(\theta) \eta(\theta) d\theta + \int_{|\theta|> \delta^*} e^{-i\theta^* \phi}(\varepsilon \theta) \hat{1}_{B_a}(\theta) f_A(\theta) d\theta = I(s) + II(s) + III(s),
\]

where \( \hat{\phi} \) denotes the Fourier transform of \( \phi \). The first summand \( I(s) \) in (4.13) can be analyzed by making use of (4.12). As in CW (p. 361), it can be shown that \( \Delta_M(\theta) \) has \( N \) derivatives with respect to \( \theta_1 \) that are integrable over \( |\theta| \leq \delta^* \), so that applying integration by parts \( N \) times yields

\[
\int_{0<|\theta|\leq \delta^*} e^{-i\theta^* \phi}(\varepsilon \theta) \hat{1}_{B_a}(\theta) \Delta_M(\theta) \nu_{\theta} Q_{\theta^*} h_A d \theta = O(s_1^{-N})
\]
as \( s_1 \to \infty \). Using a Taylor expansion of \( \nu_{\theta} Q_{\theta^*} h_A \) (which has continuous partial derivatives of order \( [r] \)), we have the following analogue of CW to represent \( I(s) \):

\[
(4.14) \quad I(s) = \sum_{j=0}^{M-2} \sum_{l=3j}^{M_j} \phi_\varepsilon * 1_{B_a} * P_{l,j}(s) + O(s_1^{-N}),
\]

where the Fourier transform of \( P_{l,j} \) is \( \pi_{l,j}(\theta) \nu_{\theta} Q_{\theta^*} h_A / (-i\theta_1 + [1/2] |\theta|^2)^j \) and \( \pi_{l,j}(\theta) \) are polynomials obtained in the multinomial expansion of \( (-\sum_{m=3}^M p_m(\theta))^j \). As in CW (pp. 362 and 365), the \( \omega_j \) in (2.7) are defined by

\[
(2\pi s_1)^{-(d-1)/2} e^{-|\bar{\pi}|^2/2s_1} s_1^{-k/2} \omega_j(s_1^{1/2}) = \sum_{(l,m):l-2m=j} p_l,m(s)
\]

for \( s_1 > 0 \). Moreover, the Fourier transform of \( (2\pi s_1)^{-(d-1)/2} 1_{[s_1 > 0]} \) is \( (-i\theta_1 + [1/2] |\theta|^2)^{-1} \); see Section 2 of Carlsson (1982).

To analyze \( III(s) \), we make use of Lemma 3 which implies that \( f_A(\theta) \) and its partial derivatives up to order \([r]\) are bounded, and \textit{a fortiori} locally integrable, in the region \( \{ \theta : |\theta| \geq \delta \} \). Moreover, \( \hat{1}_{B_a} \) and its derivatives are bounded by a constant times \( \prod_{i=1}^d |\theta_i|^{-1} \) as \( |\theta| \to \infty \). Therefore \( N \) integrations by parts as in CW (p. 359) can be used to show that \( III(s) = O(|\log \varepsilon|/s_1^{N}) \). Since \( \eta(\theta) \) and its partial derivatives of order \([r]\) are bounded for \( 0 < |\theta| \leq \delta^* \), \( N \) integrations by parts can also be used to show that \( II(s) = O(s_1^{-N}) \). The rest of the proof of Theorem 1 is essentially the same as that in CW. \( \Box \)
Proof of Theorems 2-3. Theorem 2 follows from Theorem 3 by setting \( b = s_1 \) and letting \( f \) be the indicator function of \([s_1, s_1 + h] \times \sqrt{s_1}(s_1 \gamma + C)\). To prove Theorem 3, we assume as before that \( \gamma = 0 \) and require further that \( \phi \) be a probability density function on \( \mathbb{R}^d \). Let \( \Phi_\varepsilon \) be the probability measure on \( \mathbb{R}^d \) with probability density \( \phi_\varepsilon \), and let \( u_\varepsilon \) be the Radon-Nikodym derivative of \( \Phi_\varepsilon * U_\nu^A \) with respect to Lebesgue measure on \( \mathbb{R}^d \). By Fourier inversion and an argument similar to that in the proof of Theorem 1, it can be shown that with \( \varepsilon = b^{-\eta} \) and \( \eta \) sufficiently large, \( u_\varepsilon(s) = \psi_{1,\nu}^{A,\nu}(s) + o(b^{-(1+\delta)/2}) \) uniformly in \( \tilde{s} \in [b, b + h] \) and \( \tilde{s} \in \mathbb{R}^{d-1} \). Theorem 3 then follows from the smoothing inequality

\[
| \int f d(\Phi_\varepsilon * U_\nu^A) - \int f dU_\nu^\lambda | \leq \int \Omega_f(\cdot; \varepsilon) d(\Phi_\varepsilon * U_\nu^A);
\]

see Keener (1990), p. 144. □

Remark. For the case of i.i.d. \( \xi_n \), Keener (1990) uses a considerably more elaborate method to prove a sharper version of Theorem 3, in which the moment conditions are weaker and "\( b^{-\eta} \) for every \( \eta > 0 \)" is replaced by "\( e^{-\eta b} \), for some \( \eta > 0 \)". His method involves symmetrizing the renewal measure and convolving it with \( P_0^m \) for some suitably chosen \( m \) (depending on \( b \)), where \( P_0 \) assigns probability 1/2 to the origin 0 and follows the distribution of \( \xi_1 \) with probability 1/2.

4.2. A variant of Theorem 3 and proof of Theorems 5-7

The preceding proof of Theorem 3 involves an asymptotic approximation to the Radon-Nikodym derivative of \( \Phi_\varepsilon * U_\nu^A \) with respect to Lebesgue measure on \( \mathbb{R}^d \). Suppose \( P_x \) is absolutely continuous with respect to \( M \) for every \( x \in \mathcal{X} \), as in the paragraph preceding Theorem 7 which introduces the density function \( u(x; x_0, B) \) such that

\[
U_{\nu}^A(B) = \int_A u(x; x_0, B) dM(x) \quad \text{for all } A \in \mathcal{A}.
\]

(4.15)

Note that for fixed \( x \) and \( x_0 \), \( u(x; x_0, \cdot) \) is a measure on \( \mathbb{R}^d \). Therefore a straightforward modification of the arguments used to prove Theorem 3 and Lemma 2 can be used to prove a corresponding result for \( u(x; x_0, \cdot) \).

Theorem 3*. With the same notation and assumptions as in Theorem 3, assume further that \( P_x \) is absolutely continuous with respect to \( M \) for every \( x \in \mathcal{X} \) so that (4.15) holds. Then for every \( \eta > 0 \), as \( b \to \infty \),

\[
\int f(s) u(x; x_0, ds) = \int f(s) \psi_1^{x,x_0}(s) ds + O \left( \int \Omega_f(\cdot; b^{-\eta}) \psi_1^{x,x_0}(s) ds \right) + o(b^{-(1+\delta)/2})
\]

uniformly in \( f \in \mathcal{F}_b \) and \( x, x_0 \in \mathcal{X} \).

Proof of Theorem 7. Making use of the ladder random walk, we can reduce to the special case \( P_x(\xi_{1,1} \geq 0) \) for all \( x \in \mathcal{X} \), in which \( P_+ = P \), \( \pi_+ = \pi \), \( \gamma_+ = \gamma \) and \( \mu^* = \mu \). Moreover,
for notational simplicity, we shall assume that \( d = 2 \) and \( C = (-\infty, c) \). Since \( \{T_b = n + 1\} = \{S_{n,1} \leq b, S_{n,1} + \xi_{n+1,1} > b\} \), we can write

\[
P_{x_0} \left\{ X_{T_b} \in dy, \ S_{T_b,1} > b + u, \ \sqrt{\mu_1/b}(S_{T_b,2} - b\gamma) < c \right\}
\]

\[
= \sum_{n=0}^{\infty} \int_{s_1 > u} \int_{x} P_{x_0} \left\{ X_n \in dx, \ b + u - s_1 < S_{n,1} \leq b, \ \sqrt{\mu_1/b}(S_{n,2} + s_2 - b\gamma) < c \right\} P(x, dy \times ds)
\]

\[
= \int_{x \in \mathcal{X}} \int_{s \in \mathbb{R}^d, s_1 > u} \int_{t=0}^{s_1-u} \sum_{n=0}^{\infty} P_{x_0} \left\{ X_n \in dx, \ b - S_{n,1} \in dt, \ \frac{S_{n,2} - \gamma S_{n,1}}{\sqrt{b/\mu_1}} < c + \frac{\gamma t - s_2}{\sqrt{b/\mu_1}} \right\} P(x, dy \times ds).
\]

For the rest of the proof, we make use of Theorem 3* and proceed as in the proof of Theorem 4.1 of Keener (1988), noting that \( (b + t)^{-1/2} = b^{-1/2} + O(b^{-3/2}) \) as \( b \to \infty \), uniformly in \( t \) belonging to compact sets. \( \square \)

**Proof of Theorems 5 and 6.** To prove Theorem 5, we can make use of the functional central limit theorem and the strong law of large numbers for \( u \)-uniformly ergodic Markov random walks (see Section 17.5 of Meyn and Tweedie (1993)) and thereby extend the standard arguments to prove the corresponding results for i.i.d. \( \xi_n \) (see the proof of Theorem 2 in Stam (1968)). Making use of Theorem 5, we can then prove Theorem 6 as in the case of i.i.d. \( \xi_n \) (see Lai and Siegmund (1977) for the case \( d = 1 \)). \( \square \)

### 4.3. Proof of Theorem 4

A major difference between the case \( d > 1 \) and \( d = 1 \) is that whereas the Fourier transform \( f_A \) in Lemma 2 is integrable in a neighborhood of 0 for \( d > 1 \) (e.g., \( \int_{S_2}^{\infty} f_{S_2}^{\infty} d s_1 d s_2 = 4 \pi \sqrt{\delta} \) for \( d = 2 \)), such local integrability does not hold for \( d = 1 \). For \( d > 1 \), because of this local integrability, although Lemma 2(ii) only represents \( f_A(\theta) \) for \( 0 < \theta \leq \delta^* \), we can ignore how \( f_A \) is defined at 0. The situation is dramatically different for the case \( d = 1 \), for which \( f_A(0) \) becomes a delta function. Specifically, for \( |\theta| \leq \delta^* \),

\[
(4.16) \quad f_A(\theta) = \left\{ (1 - \lambda(\theta))^{-1} \nu Q h_A + \eta(\theta) \right\} 1_{\{0 < |\theta| \leq \delta^*\}} + \pi(A)(\pi/\mu) \delta(\theta),
\]

where \( \delta(\theta) \) is the delta function. This can be proved by an argument similar to that of Carlsson (1983, p. 147), which involves summing \( E_{v}(e^{i \theta \sum_{n} h_A(X_n)}) \) over \( 0 \leq n \leq N-1 \) and showing that \( \lambda^N(\theta) \{ (1 - \lambda(\theta))^{-1} - (-i \mu \theta)^{-1} \} \) and \( \psi_N(\theta) := E_{v}(e^{i \theta \sum_{n} h_A(X_n)}) \{ (i \theta)^{-1} + \pi \delta(\theta) \} \) converge to 0 in the space \( T' \) of tempered distributions. In this connection, note that \( (i \theta)^{-1} + \pi \delta(\theta) \) is the Fourier transform of the Heaviside function \( H(x) = 1_{\{x \geq 0\}} \) and therefore \( \psi_N \) is the Fourier transform of \( F_N(x) = P_{v}(S_n \leq x, X_n \in A) \). Since \( 1 - \lambda(\theta) \sim -i \mu \theta \), an argument similar to that of Stone (1965, p. 331) can be used to show that the real part \( R(1 - \lambda(\theta)) \) is locally integrable. Therefore, we can proceed as in the proof of Theorem 1 to show that \( (U_{v}^{A} - m_{A})([b, b + h]) = o(b^{-(r-1)}) \) as \( b \to \infty \), where \( dm_{A}(x) = \mu^{-1} \pi(A) ds \). \( \square \)
Remark. For the case of i.i.d. \( \xi_n \), Stone (1965) and Carlsson (1983) have also derived asymptotic formulas for \( U([0, b]) \) and thereby also for \( U((\infty, b]) \). In particular, if the \( \xi_n \) are nonnegative and strongly nonlattice, then \( \lim_{b \to \infty} \frac{b}{\mu} \to \frac{\sigma^2 + \mu^2}{2\mu^2} \). Note, however, that \( U([0, b]) = ET_b \) for nonnegative \( \xi_n \) and \( E\xi T_b - b/\mu \) converges to a more complicated limit as \( b \to \infty \) for Markov-dependent \( \xi_n \), as shown in Theorem 8.

Appendix: Characteristic functions of Markov random walks

Using the same notation and assumptions as in the first paragraph of Section 2, define \( P_\theta, P, \nu_\theta \) and \( Q \) by (4.2)-(4.4). Condition (2.2) ensures that \( P_\theta \) and \( P \) are bounded linear operators on \( B \), and (2.1) implies that

\[
\|P^n - Q\| = \sup_{h \in B : \|h\|_w = 1} \|P^n h - Qh\|_w \leq \gamma \rho^n. \tag{A.1}
\]

For a bounded linear operator \( T : B \to B \), the resolvent set is defined as \( \{z \in \mathbb{C} : (T - zI)^{-1} \text{ exists}\} \) and \( (T - zI)^{-1} \) is called the resolvent (when the inverse exists). From (A.1) it follows that for \( z \neq 1 \) and \( |z| > \rho \),

\[
R(z) := \frac{Q}{(z - 1)} + \sum_{n=0}^{\infty} (P^n - Q)/z^{n+1} \tag{A.2}
\]
is well defined. Since \( R(z)(P - zI) = -I = (P - zI)R(z) \), the resolvent of \( P \) is \( -R(z) \). Moreover, by (2.3) and an argument similar to the proof of Lemma 2.2 of Jensen (1987), there exist \( K > 0 \) and \( \eta > 0 \) such that for \( |\theta| \leq \eta \), \( |z - 1| > (1 - \rho)/6 \) and \( |z| > \rho + (1 - \rho)/6 \), \( \|P_\theta - P\| \leq K|\theta| \) and \( R_\theta(z) := \sum_{n=0}^{\infty} R(z)(P_\theta - P)R(z)^n \) is well defined. Since \( R_\theta(z)(P_\theta - zI) = R_\theta(z)(P_\theta - P) + (P - zI) = -I = (P_\theta - zI)R_\theta(z) \), the resolvent of \( P_\theta \) is \( -R_\theta(z) \).

For \( |\theta| \leq \eta \), the spectrum (which is the complement of the resolvent set) of \( P_\theta \) therefore lies inside the two circles \( C_1 = \{z : |z - 1| = (1 - \rho)/3\} \) and \( C_2 = \{z : |z| = \rho + (1 - \rho)/3\} \). Hence by the spectral decomposition theorem (cf. Riesz and Sz-Nagy (1955), p. 421), \( B = B_1(\theta) \oplus B_2(\theta) \) and

\[
Q_\theta := \frac{1}{2\pi i} \int_{C_1} R_\theta(z)dz, \quad I - Q_\theta := \frac{1}{2\pi i} \int_{C_2} R_\theta(z)dz \tag{A.3}
\]
are parallel projections of \( B \) onto the subspaces \( B_1(\theta), B_2(\theta) \) respectively. Moreover, by an argument similar to the proof of Lemma 2.3 of Jensen (1987), there exists \( 0 < \delta \leq \eta \) such that \( B_1(\theta) \) is one-dimension for \( |\theta| \leq \delta \) and \( \sup_{|\theta| \leq \delta} \|Q_\theta - Q\| < 1 \). For \( |\theta| \leq \delta \), let \( \lambda(\theta) \) be the eigenvalue of \( P_\theta \) with corresponding eigenspace \( B_1(\theta) \). Since \( Q_\theta \) is the parallel projection onto the subspace \( B_1(\theta) \) in the direction of \( B_2(\theta) \), (4.5) holds. Therefore, for \( h \in B \),

\[
E_{\nu}\{e^{it\theta S_n} h(X_n)\} = \nu \theta P_\theta^n h = \nu \theta P_\theta^n (Q_\theta + (I - Q_\theta))h = \lambda^n(\theta)\nu \theta Q_\theta h + \nu \theta P_\theta^n (I - Q_\theta)h.
\]
Suppose (2.9) also holds. An argument similar to the proof of Lemma 2.4 of Jensen (1987) shows that there exist $0 < \delta^* < \delta$ and $K^* > 0$ such that for $|\theta| \leq \delta^*$, $|\nu_\theta P_\theta^n (I - Q_\theta) h| \leq K^* \|h\|_\infty |\theta| \{(1 + 2\rho)/3\}^n$. Moreover, analogous to Lemmas 2.5, 2.6 and 2.7 of Jensen (1987), it can be shown that $\lambda(\theta)$, $\nu_\theta Q_\theta h$ and $\sum_{n=0}^\infty \nu_\theta P_\theta^n (I - Q_\theta) h$ have continuous partial derivatives of order $[r]$ for $|\theta| \leq \delta^*$.

Acknowledgment

The research of Cheng-Der Fuh was supported by the National Science Council of the Republic of China. The research of Tze Leung Lai was supported by the National Science Foundation, the National Security Agency, and the Center for Advanced Study in the Behavioral Sciences.

References


