SMALL DRIFT LIMIT THEOREMS FOR RANDOM WALKS

ERNST SCHULTE-GEERS AND WOLFGANG STADJE

Abstract. We show analogs of the classical arcsine theorem for the occupation time of a random walk in \((−∞, 0)\) in the case of a small positive drift. To study the asymptotic behavior of the total time spent in \((−∞, 0)\) we consider parametrized classes of random walks, where the convergence of the parameter to zero implies the convergence of the drift to zero. We begin with shift families, generated by a centered random walk by adding to each step a shift constant \(a > 0\) and then letting \(a\) tend to zero. Then we study families of associated distributions. In all cases we arrive at the same limiting distribution, which is the distribution of the time spent below zero of a standard Brownian motion with drift 1. For shift families this is explained by a functional limit theorem. Using fluctuation-theoretic formulas we derive the generating function of the occupation time in closed form, which provides an alternative approach. In the course also give a new form of the first arcsine law for the Brownian motion with drift.

2010 Mathematics Subject Classification: 60G50, 60F05.

1. Introduction

For the classical symmetric random walk with \(±1\) steps it is well known that the three random variables “time spent on the positive axis”, “position of the first maximum” and “last exit from zero” are identically distributed and (suitably normalized) asymptotically arcsine-distributed. Here the norming factor is the length of the time interval the random walk has been observed, so that the limiting statements refer to “relative” times.

Consider now a classical random walk with drift \(δ \neq 0\). Clearly the same “relative” variables can be studied. The asymptotic distribution of the random variable “(fraction of) time spent in \((−∞, α]\) has been determined by Takács [21], by applying a functional limit theorem. But if \(δ \neq 0\) there is also another, “absolute” perspective. If for example \(δ > 0\) for a general random walk, it is clear that \(Z(δ) = \text{“number of visits in } (−∞, 0) \text{” is almost surely finite, and that } Z(δ) \to \infty \text{ in probability as } δ \searrow 0\). One may ask if \(Z(δ)\), after multiplication with some deterministic function \(a(δ)\), has a non-degenerate limit distribution. This paper aims to answer these and related questions for random walks in the heavy-traffic regime, i.e., when the drift converges to zero. In all cases the limiting distribution for the occupation time in \((−∞, 0)\), properly rescaled, turns out to have the density

\[
p(t) = 2 \frac{\varphi(\sqrt{2t})}{\sqrt{2t}} - 2 \Phi(-\sqrt{2t}), \quad t > 0
\]

where \(\varphi\) and \(\Phi\) are the density and the distribution function of \(N(0, 1)\), respectively.

The distribution of the occupation time in \((−∞, 0)\) of Brownian motion with positive drift also has density \(p(t)\), and in Section 2 we begin with related results for Brownian motion.

Key words and phrases. Random walk; transient; occupation time; arcsine law; small drift; limit distribution.

1

arXiv:1601.00609v2 [math.PR] 28 May 2016
We show for example that the distribution of the time of the last exit from 0 of Brownian motion with drift during a finite time interval is composed of the arcsine and a truncated exponential distribution. In Section 2 we derive the limiting occupation time distribution for shift families generated from a centered random walk by adding to each step a shift constant \( a > 0 \) and then letting \( a \) tend to zero. The proof that (1.1) gives the asymptotic distribution is based on Donsker’s invariance principle. In Section 4 we give the key fluctuation-theoretic formulas for the distribution for the occupation time in \((-\infty, 0)\) for general random walks.

The arcsine law and its ramifications are a classical topic but there are always recent contributions, for example some new explicit distributions [15], new proofs [9], or asymptotic considerations [16]. Interesting results on the number of visits to one point by skipfree random walks and related questions can be found in [4]. The problem considered in this paper is also connected to the heavy traffic approximation problem in queueing theory, in which the growth of the all-time maximum of \( S_n - na \) (where \( S_n \) is the \( n \)th partial sum of iid random variables with mean zero) is studied as \( a \downarrow 0 \). In the queueing context this is equivalent to the growth of the steady-state waiting time in a \( GI/G/1 \) system when the traffic load tends to 1. This question was first posed by Kingman (see [13]) and was investigated by many authors (e.g. [3, 14, 17, 18, 20]).

2. Occupation times and last exit from 0 for Brownian motion with drift

We start by presenting two results on occupation times for Brownian motion with positive drift \( \delta > 0 \) and variance \( \sigma^2 \), one known and one new. Let \( B_t \) be a standard Brownian motion and \( X_t = \sigma B_t + \delta t \).

**Lemma 2.1.** (1) Let \( z > 0 \) and \( T_z = \inf \{ t \geq 0 : X_t \geq z \} \) be the first time when \( X_t \) reaches level \( z \). Then \( T_z \) has Laplace transform
\[
\ell_{T_z}(s) = E e^{-sT_z} = \exp \left( -\frac{z}{\sigma^2} \left( \sqrt{\delta^2 + 2\sigma^2 s} - \delta \right) \right).
\]

(2) Let \( V_0 = V_0(\delta) = \int_0^\infty 1_{(-\infty, 0)}(X_t) \, dt \) be the total time that \( X_t \) spends below zero. Then \( V_0 \) has Laplace transform
\[
\ell_{V_0}(s) = E e^{-sV_0} = \frac{2\delta}{\sqrt{\delta^2 + 2\sigma^2 s}}.
\]

Proofs for (1) resp. (2) (for \( \sigma^2 = 1 \)) can be found in [12] resp. [10]. Note \( (\delta^2/2\sigma^2)V_0 \) has the Laplace transform \( 2/(1 + \sqrt{1 + s}) \). We call \( A \) a generic random variable with this Laplace transform.

The density of \( A \) is given by (1.1). To see this, note that \( 1/\sqrt{1 + s} \) is the Laplace transform of the gamma distribution \( \Gamma_{1, \frac{1}{2}} \), which has density
\[
\gamma_{1, \frac{1}{2}}(t) = 1_{(0, \infty)}(t) \frac{e^{-t}}{\sqrt{\pi t}}.
\]

Therefore \( [1 - (1/\sqrt{1 + s})]/s \) is the Laplace transform of \( 1 - \Gamma_{1, \frac{1}{2}}(t) = \int_t^\infty \gamma_{1, \frac{1}{2}}(x) \, dx \). The equality
\[
\frac{1}{1 + \sqrt{1 + s}} = \frac{1}{\sqrt{1 + s}} - \frac{1}{s} \left( 1 - \frac{1}{\sqrt{1 + s}} \right)
\]
now yields density (1.1).
For $z \geq 0$ let $V_z = \int_0^\infty 1_{(-\infty,z)}(X_t) \, dt$ the total time the process spends below $z$. Then the obvious decomposition (obtained by conditioning on $T_z$) $V_z = T_z + V'_0$ (where $V'_0$ is independent of $T_z$ and distributed as $V_0$) yields

**Lemma 2.2.** $V_z$ has Laplace transform

$$\ell_{V_z}(s) = E(e^{-sV_z}) = \ell_{T_z}(s) \ell_{V_0}(s).$$

The density and distribution function are given in [10].

We focus in the sequel on the time spent on the negative axis, but it is also of interest to look at the other classical arcsine variable, i.e., the time of the last exit from 0. Here we determine its distribution. Let $\delta \in \mathbb{R} \setminus \{0\}, \sigma^2 = 1$, so that $X_t = B_t + \delta t$, and consider $W = \sup\{t \in [0,1] : X_t = 0\}$, the last time $X_t$ visits 0 in $[0,1]$.

Recall that for $\delta = 0$, i.e., for the standard Brownian motion, the standard arcsine distribution (which has density $1_{(0,1)}(t)(1/\sqrt{t(1-t)})$ and distribution function $(2/\pi) \arcsin(\sqrt{t})$ on $[0,1]$) is the distribution of the last exit time from zero in the interval $[0,1]$.

The distribution of $W$ turns out to have a nice representation in terms of the standard arcsine distribution and a truncated exponential distribution. As this result seems new, we provide a proof.

**Theorem 2.3.** $W \overset{\text{d}}{=} C \cdot \min\{1, D_\delta\}$ where $C$ and $D_\delta$ are independent, $C$ is arcsine-distributed, and $D_\delta$ is $\exp(\delta^2/2)$-distributed. The moments of $W$ are given by

$$E W^k = \binom{2k}{k} \frac{1}{2^{2k}} \int_0^1 ky^{k-1} e^{-\delta^2 y/2} \, dy, \quad k \geq 1.$$  

**Proof.** We use a random walk approximation in the style of Takács [21]. Let $Y_1, Y_2, \ldots$ be iid with

$$P(Y_i = 1) = p = \frac{1}{2} + \frac{\delta}{2\sqrt{n}}, \quad P(Y_i = -1) = q = 1 - p$$

($p$ and $q$ depend on $n$, but this is suppressed in the notation) and partial sums $S_0 = 0, S_k = \sum_{i=1}^k Y_i$.

It is easy to see that the processes $X^{(n)}$ defined by

$$X^{(n)}(t) = \frac{1}{\sqrt{n}} S_{\lfloor nt \rfloor}, \quad 0 \leq t \leq 1$$

converge in distribution to $X = (X_t)_{t \in [0,1]}$ in $D[0,1]$.

Furthermore, the last-exit time from 0 is continuous in the Skorohod topology on $D[0,1]$ on a set of $P_X$-measure 1, and

$$T_n = \sup\{t \in [0,1] : X^{(n)}(t) = 0\} = \frac{1}{n} \max\{0 \leq k \leq n : S_k = 0\} =: M_n/n$$

Then it suffices to show that $M_N/N \longrightarrow C \cdot \min\{1, D_\delta\}$ as $N \longrightarrow \infty$.

Since $1/\sqrt{1 - 4pqz^2}$ and $(\sqrt{1 - 4pqz^2})/(1 - z)$ are the generating functions for the sequences of probabilities $P(S_n = 0)$ and $P(S_1 \neq 0, \ldots, S_n \neq 0)$, respectively, the generating function of
\[ E_t^{M_N} = \sum_{k=0}^{N} t^k P(S_k = 0, S_{k+1} \neq 0, \ldots, S_N \neq 0) \]
\[ = \sum_{k=0}^{N} t^k P(S_k = 0) P(S_1 \neq 0, \ldots, S_{N-k} \neq 0) \]
\[ = [z^N] \frac{1}{\sqrt{1 - 4pqz^2}} \frac{\sqrt{1 - 4pqz^2}}{1 - z} \]
\[ = [z^N] \frac{1}{\sqrt{1 - 4pqz^2}} \frac{\sqrt{1 - 4pqz^2}}{1 - z^2} (1 + z). \]

(Here and in the following \([z^N]f(z)\) denotes the coefficient of \([z^N]\) in the Taylor expansion of the function \(f(z)\) around zero.) Thus the generating functions for \(N = 2n + 1\) and \(N = 2n\) are identical and it is enough to consider even \(N\). Let \(N = 2n\) be even (and \(n > \delta^2/2n\)) and \(U_n = M_N/2\). Then the generating function of \(U_n\) is
\[ E_t^{U_n} = [z^{2n}] \frac{1}{\sqrt{1 - 4pqz^2}} \frac{\sqrt{1 - 4pqz^2}}{1 - z^2} \]
so that the \(k\)-th factorial moment \(u_{k,n} = E(U_n(U_n - 1) \cdots (U_n - k + 1))\) of \(U_n\) is given by
\[ u_{k,n} = k! (-1)^k \left( \frac{-1}{k} \right) (4pq)^k [z^{n-k}] \frac{1}{(1-4pqz)^k (1-z)} \]
\[ = k(-1)^k \left( \frac{-1}{k} \right) (4pq)^k [z^{n-k}] \frac{1}{(1-z)} \int_0^\infty x^{k-1} e^{-(1-4pq)x} \, dx \]
\[ = (-1)^k \left( \frac{-1}{k} \right) (4pq)^k \int_0^\infty kx^{k-1} e^{-x} \left( \sum_{j=0}^{n-k} \frac{(4pqx)^j}{j!} \right) \, dx. \]

Now denote by \(Poiss(\lambda)\) a random variable having the Poisson distribution with parameter \(\lambda\). As \(4pq = 1 - (\delta^2/2n)\), we obtain
\[ \int_0^\infty kx^{k-1} e^{-x} \left( \sum_{j=0}^{n-k} \frac{(4pqx)^j}{j!} \right) \, dx \]
\[ = \int_0^\infty kx^{k-1} e^{-x(1-4pq)} \, P \left( Poiss(4pqx) \leq n - k \right) \, dx \]
\[ = n^k \int_0^\infty ky^{k-1} e^{-\delta^2 y/2} \, P \left( Poiss((n - \delta^2/2)y) \leq n - k \right) \, dy. \]

By the central limit theorem,
\[ P \left( Poiss((n - \delta^2/2)y) \leq n - k \right) \rightarrow \begin{cases} 
1 & \text{for } 0 \leq y < 1 \\
\frac{1}{2} & \text{for } y = 1 \\
0 & \text{for } y > 1 
\end{cases} \]
so that for every $k$ we have

$$
\frac{u_{n,k}}{n^k} \rightarrow (-1)^k \left( \frac{1}{2} \right) \int_0^1 ky^{k-1} e^{-y\delta^2/2} \, dy.
$$

Hence $\mathbb{E} T_N^k / N^k$ tends to the same limit. This shows the second assertion. Finally,

$$
\mathbb{E} C^k = (-1)^k \left( \frac{1}{2} \right) \left( \frac{2k}{k} \right) \frac{1}{2^{2k}}
$$

and integration by parts shows that $\int_0^1 ky^{k-1} e^{-y\delta^2/2} \, dy = \mathbb{E} \min\{1, D_\delta^k\}$. Thus all moments of $T_N / N$ converge to the corresponding moments of $C \cdot \min\{1, D_\delta\}$. Since the distribution of $C \cdot \min\{1, D_\delta\}$ is clearly determined by its moments the first assertion follows.

**Remark 2.4.** As an immediate consequence of the scaling properties of Brownian motion we see that the distribution of

$$
W_T = \sup\{t \leq T : \sigma B_t + \delta t = 0\}
$$

is the same as that of $C \cdot \min\{T, D_\delta/\sigma\}$. The time of the last zero of $\sigma B_t + \delta t$ in the interval $[0, \infty)$ is thus distributed as $C \cdot D_\delta/\sigma$, which is the gamma distribution with parameters $\delta^2/2\sigma^2$ and 1/2.

**Remark 2.5.** Clearly $V_0$ (the occupation time on the negative axis) is stochastically smaller than $W_\infty$ (the last exit time from zero), and the results above quantify this precisely. We find e.g. that

$$
\mathbb{E}(V_0) = \frac{\sigma^2}{2\delta^2} = \frac{1}{2} \mathbb{E}(W_\infty).
$$

**Remark 2.6.** Last-exit times of Brownian motion from moving boundaries have been studied intensively, and more complicated expressions for the density of the last-exit time from a linear boundary were derived in [19] and [11]. The representation in (2.3) appears to be new, as it is not mentioned in the encyclopedic monograph [2]. For the density of the sojourn time found by Takács by a random walk limit two “purely Brownian” explanations have been given in [6]. It is natural to ask for such an explanation for the representation in (2.3).

3. Limit of occupation times for shifted random walks

In this section we consider a shifted random walk. Specifically, let $(X_{\delta_1}, X_{\delta_2}, \ldots)$ be a parametrized sequence of iid random variables with $\mathbb{E}(X_{\delta,i}) = 0$, $\text{Var}(X_{\delta,i}) = \sigma^2(\delta) \in (0, \infty)$. Let $\delta > 0$ and

$$
Y_{\delta,i} = X_{\delta,i} + \delta, \quad S_{\delta,n} = \sum_{i=1}^n X_{\delta,i}, \quad S_n^\delta = \sum_{i=1}^n Y_{\delta,i}.
$$

We are interested in the occupation time

$$
Z_0^\delta = \sum_{i=1}^\infty 1_{(-\infty,0)}(S_n^\delta).
$$
Throughout this section we assume that $\sigma^2(\delta) \to \sigma^2 > 0$ as $\delta \to 0$ and that the following Lindeberg-type condition holds: for every $\varepsilon > 0$,

$$
\lim_{\delta \to 0} \int_{|\delta X_{\delta,1}| > \varepsilon} X_{\delta,1}^2 \, dP = 0. \tag{3.1}
$$

These conditions are chosen such that for the triangular array with the variables

$$
Z_{\delta,k} = \frac{\delta}{\sigma(\delta)} X_{\delta,k}, \quad k = 1, \ldots, \left\lfloor \frac{1}{\delta^2} \right\rfloor
$$

the central limit theorem holds: indeed, $\text{Var}(Z_{\delta,1}) = \delta^2$ and the Lindeberg condition for this triangular array reads as

$$
\lim_{\delta \to 0} \frac{1}{\delta^2} \int_{|Z_{\delta,1}| > \varepsilon \delta^2 \left\lfloor \frac{1}{\delta} \right\rfloor} Z_{\delta,1}^2 \, dP = \lim_{\delta \to 0} \frac{1}{\delta^2} \sigma^2(\delta) \int_{|\delta X_{\delta,1}| > \varepsilon \sigma(\delta) \delta^2 \left\lfloor \frac{1}{\delta^2} \right\rfloor} X_{\delta,1}^2 \, dP = 0 \quad \text{for every} \quad \varepsilon > 0,
$$

which is clearly true under the conditions above.

We use similar ideas as Prohorov [17], who proved the following:

**Theorem 3.1. (Prohorov)** In the situation above let $M^\delta = \min\{S_{\delta,n}^k : n \geq 0\}$. Then

$$
P(\delta M^\delta > x) \to e^{-2x/\sigma^2} \quad \text{for all} \quad x > 0.
$$

In [17] the maximum in the case of negative drift was considered instead of $M^\delta$. The result had been proved earlier by Kingman under the assumption of the existence of an exponential moment.

The following lemma will be needed to obtain tightness bounds.

**Lemma 3.2.** In the situation above let $z \geq 0$ and let $\delta_k > 0$ be a sequence of positive numbers satisfying $\sup_{k \geq 1} \sigma^2(\delta_k) < \infty$. Then for every $\varepsilon > 0$ we can find a $T$ such that for all $k$

$$
P\left( \sup_{n \geq T/\delta_k^2} (|S_{\delta_k,n} - n\delta_k| - z/\delta_k) \geq -\varepsilon \right) < \varepsilon.
$$

**Proof.** First consider a sequence $S_n$ of partial sums of an arbitrary iid sequence $(X_i)$ with $E(X_1) = 0$ and $\text{Var}(X_1) = \sigma^2$. Let $a, b > 0$, $Na > b$ and consider the event $E_N = \{\sup_{n \geq N} (|S_n| - na) \geq -b\}$. Clearly

$$
E_N = \bigcup_{j=0}^{\infty} \left\{ \max_{2^j N \leq n < 2^{j+1} N} (|S_n| - na) \geq -b \right\}
$$

$$
\subseteq \bigcup_{j=0}^{\infty} \left\{ \max_{2^j N \leq n < 2^{j+1} N} |S_n| \geq 2^j Na - b \right\}
$$

$$
\subseteq \bigcup_{j=0}^{\infty} \left\{ \max_{n \leq 2^j + 1} |S_n| \geq 2^j Na - b \right\}.
$$

By Kolmogorov’s inequality,

$$
P\left( \max_{n \leq 2^j + 1} |S_n| \geq 2^j Na - b \right) \leq \frac{2^{j+1} N \sigma^2}{(2^j Na - b)^2}.
$$
Now set \( N = \frac{T}{\delta}, a = \delta, b = \frac{z}{\delta}, X_i = X_{\delta i}. \) It follows that
\[
\mathbb{P}(\sup_{n \geq T/\delta^2} (|S_{\delta n} - n\delta| \geq \frac{z}{\delta}) \leq \sum_{j=0}^{\infty} \frac{2^{j+1}T\sigma^2(\delta)}{(2^j T - z)^2}.
\]
The bound on the right side depends on \( \delta \) only via \( \sigma^2(\delta) \) and can clearly be made arbitrarily small (under the assumptions above).

**Corollary 3.3.** In the situation above let \( z \geq 0 \) and let \( \delta_k > 0 \) be a sequence of positive numbers satisfying \( \sup_{k \geq 1} \sigma^2(\delta_k) < \infty. \) Then for every \( \epsilon > 0 \) one can find a \( T \) such that for all \( k \)
\[
\mathbb{P}(\min_{n \geq T/\delta^2_k} \delta_k(S_{\delta_k n} + n\delta_k) \leq z) < \epsilon.
\]

**Theorem 3.4.**
\[
\frac{\delta^2}{2\sigma^2(\delta)} Z_0^\delta \rightarrow A \text{ in distribution as } \delta \searrow 0.
\]

**Proof.** By the remark following 2.1 it suffices to show that \( \delta^2 Z_0^\delta \rightarrow V_0 \) in distribution, where \( V_0 \) is the distribution of the time the process \( W_t = \sigma B_t + t \) spends below zero.

Let \( T > 0 \) and consider the sequence of processes
\[
U^\delta(t) = \delta \sum_{i=1}^{\lfloor t/\delta^2 \rfloor} Y_i^\delta, \quad 0 \leq t \leq T.
\]
By Donsker’s limit theorem (in the version for triangular arrays, see e.g. [1], p.147), the sequence \( U^\delta \rightarrow \sigma B + id \) in distribution in \( D[0, T] \), where \( \sigma B + id \) denotes the Brownian motion with variance \( \sigma^2 \) and drift 1, i.e., with coordinate variables \( \sigma B_t + t \). For any bounded Borel function \( v \) on \( [0, T] \) the functional \( x \mapsto \int_0^T v(x_t) \, dt \) on \( D[0, T] \) is Skorohod-measurable and continuous except on a set of \( B \)-measure 0 (see e.g. [1], p. 247). Thus,
\[
\delta^2 \text{card}(\{n : S_n^\delta < 0, 1 \leq n \leq T/\delta^2\}) = \int_0^{\delta^2 T/\delta^2} 1_{(-\infty, 0)}(U^\delta(t)) \, dt
\]
\[
\rightarrow \int_0^T 1_{(-\infty, 0)}(X_t) \, dt \text{ as } \delta \searrow 0
\]
in distribution and we will be done if we can justify the interchange of the limits \( T \rightarrow \infty \) and \( \delta \searrow 0 \). Let \( \delta_k > 0 \) be a sequence decreasing to zero and let \( \epsilon > 0 \). By corollary 3.3 we can find an \( N \) such that \( \mathbb{P}(\min_{n \geq N/\delta^2_k} S_n^\delta_k \leq 0) < \epsilon \) for all \( k \).

Thus,
\[
\lim_{T \rightarrow \infty} \sup_{k \geq 1} \mathbb{P}(\min_{n \geq T/\delta^2_k} S_n^\delta_k \leq 0) = 0 \quad (3.2)
\]
and the assertion follows since, by the monotone convergence theorem,
\[
\lim_{T \rightarrow \infty} \int_0^T 1_{(-\infty, 0)}(X_t) \, dt = \int_0^\infty 1_{(-\infty, 0)}(X_t) \, dt.
\]

**Remark 3.5.** A related discussion can be found in [20]. In that paper, Shneer and Wachtel derived an extension of Kolmogorov’s inequality and treated the maximum of random walks with negative drift and step size distributions attracted to a stable law of index \( \alpha \in (1, 2] \).
In the case of finite variance ($\alpha = 2$) they already remarked that their results (including in particular the crucial relation (3.2)) remain valid if the conditions assumed above hold. 

**Remark 3.6.** Assume that the $X_i$ are independent with $E(X_i) = 0$ and variances $\text{Var}(X_i) = \sigma_i^2$ and satisfy Lindeberg’s condition. Let $s^2 = \sum_{k=1}^{\infty} \sigma_k^2$. Then the step processes $X_n(t)$ which jump to the value $S_n/s_n$ at time $s_n^2/s_n^2$ converge weakly to a standard Brownian in $D[0,1]$ (by Prohorov’s extension of Donsker’s theorem). One may thus expect that they exhibit a similar limiting behavior.

Finally, replacing 0 by $z/\delta$ and repeating the steps in the proof of 3.4 yields

**Theorem 3.7.** In the situation above let $z > 0$ and $Z^\delta_z = \sum_{n=1}^{\infty} 1_{(-\infty,z)}(S^\delta_n)$. Then $\delta^2 Z^\delta_z \rightarrow V_z$ in distribution, where the Laplace transform of $V_z$ is given in Lemma 2.2 with $\delta = 1$.

If here $z$ depends on $\delta$ such that $\delta z(\delta) \rightarrow 0$ as $\delta \rightarrow 0$ we find

**Proposition 3.8.** In the situation above let $(z(\delta))$ a sequence of positive numbers with $z(\delta) = o(1/\delta)$ and $\sup_\delta z(\delta) < \infty$. Then

$$\delta^2 Z^\delta_{z(\delta)} \rightarrow V_0 = 2\sigma^2 A$$

as $\delta \rightarrow 0$.

**Proof.** Clearly $V_0$ is stochastically smaller than any distributional limit of $\delta^2 Z^\delta_{z(\delta)}$ (because $Z^\delta_0$ is stochastically smaller than $Z^\delta_y$ for $y \geq 0$), furthermore $V_y = T_y + V_0$ is stochastically smaller than $V_z$ for $y \leq z$. Let $\varepsilon > 0$ and $C = \sup_\delta z(\delta)$, then $C < \infty$ and $\delta^2 Z^\delta_{zC/\delta} \rightarrow V_{zC}$ in distribution as $\delta \rightarrow 0$ (by theorem 3.7). Since $Z^\delta_{z(\delta)} = Z^\delta_{z(\delta)/\varepsilon}$ is stochastically smaller than $Z^\delta_{zC/\delta}$ for $\delta \leq \varepsilon$, any distributional limit of $\delta^2 Z^\delta_{z(\delta)}$ is stochastically smaller than $V_{zC}$. Thus the distributional limit exists and equals $V_0$.

We close this section with an application of Theorem 3.4 in a frequently encountered situation.

**Example 3.9.** (Expectation shift in exponential families.)

Let $U$ be a non-constant real random variable such that the moment generating function

$$m(s) = Ee^{sU}$$

is finite in an open interval $I$ around 0, and $E(U) = m'(0) = 0$, $\text{Var}(Y) = \sigma^2$.

For $p \in I$ let $U_p$ have the “associated” distribution with moment generating function $m_p(s) = m(p+s)/m(p)$, clearly $U_p$ has expectation $E(U_p) = m'(p)/m(p)$ and variance $\sigma^2(p) = m''(p)m(p) - (m'(p))^2/m(p)^2$.

Let $Z_0(p)$ denote the random variable “time spent in $(-\infty,0)$” by the random walk generated by iid variables with distribution $U_p$. Then

$$\frac{(E(U_p))^2}{2\sigma^2(p)} Z_0(p) \rightarrow A$$

in distribution for $p \searrow 0$.

**Proof.** It is well known that $s \mapsto \log m(s)$ is strictly convex on $I$, thus $p \mapsto \frac{m'(p)}{m(p)} = E(U_p)$ is strictly increasing. Thus we may parameterize the distributions by $\delta(p) = E(U_p)$. We have $\delta(p) \searrow 0$ for $p \searrow 0$ and $\sigma^2(p) \rightarrow \sigma^2$ as $p \searrow 0$. Let $X_{\delta(p)} = U_p - E(U_p)$ and
\[ Y_{\delta(p)} = X_{\delta(p)} + \delta(p) = U_p. \]

Then the Lindeberg condition (3.1) is satisfied, since by Chebyshev’s inequality

\[ \int_{|\delta(p)X_{\delta(p)}| > \varepsilon} X_{\delta(p)}^2 \, dP \leq \frac{\delta^2(p)\sigma^2(p)}{\varepsilon^2} \]

and the claim follows from Theorem 3.4.

\section{The fluctuation theoretic approach}

The topics investigated here belong to the fluctuation theory of random walks. We recall some basic facts, which will be used in the sequel and can e.g. be found in Section XII.7 of [8].

We consider a random walk \((S_n)_{n \geq 1}\), i.e., a sequence of partial sums of iid random variables and let \(R = \inf\{n \geq 1 : S_n < 0\}\) and \(W = \inf\{n \geq 1 : S_n \geq 0\}\) be the lengths of the first strictly descending and weakly ascending ladder epochs of the random walk, respectively. We denote by \(r(z)\) and \(a(z)\) denote the corresponding probability generating functions and set \(\mu = EW\). The occupation time of interest is \(Z_0 = \sum_{n=1}^{\infty} 1_{(-\infty,0)}(S_n)\).

**Theorem 4.1. (Sparre Andersen)** For \(|z| < 1\)

\[
\frac{1}{1 - r(z)} = \exp \left\{ \sum_{n=1}^{\infty} \frac{z^n}{n} P(S_n < 0) \right\}
\]

\[
\frac{1}{1 - a(z)} = \exp \left\{ \sum_{n=1}^{\infty} \frac{z^n}{n} P(S_n \geq 0) \right\}
\]

An immediate consequence is the factorization theorem.

**Theorem 4.2. (Duality)** For \(|z| < 1\)

\[(1 - r(z))(1 - a(z)) = 1 - z.\]

It follows from the factorization theorem is that \(W(R)\) has a finite expected value if and only if \(R(W)\) is defective, and that the relations \(E(R)P(W = \infty) = 1\) and \(E(W)P(R = \infty) = 1\) hold.

At the combinatorial heart of fluctuation theory is the “Sparre Andersen transformation” (made explicit by Feller and refined by Bizley and Joseph) given in Lemma 3 of XII.8 of [8]:

**Lemma 4.3.** Let \(x_1, \ldots, x_n\) be real numbers with exactly \(k \geq 0\) negative partial sums \(s_{i_1}, \ldots, s_{i_k}\), where \(i_1 > \ldots > i_k\). Write down \(x_{i_1}, \ldots, x_{i_k}\) followed by the remaining \(x_i\) in their original order. (If \(k = 0\), the sequence remains unchanged). The transformation thus defined is invertible, and the first (absolute) minimum of the partial sums of the new arrangement occurs at the \(k\)-th place.

Clearly this extends to infinite sequences with exactly \(k\) negative partial sums: just apply the bijection above to an initial section large enough to contain all the negative partial sums, and leave the rest unchanged.

The following formulas express the generating function of \(Z_0\) in terms of \(r(z)\) or of \(a(z)\), respectively.
Theorem 4.4.

\[
E_z Z_0 = \frac{1 - r(1)}{1 - r(z)} = \frac{1}{\mu} \frac{1 - a(z)}{1 - z} = \exp \left\{ - \sum_{k=1}^{\infty} \frac{(1 - z^k) P(S_k < 0)}{k} \right\}.
\]  

(4.1)

Proof. According to Lemma 2.3, each sequence \( x_1, x_2 \ldots \) with exactly \( k \) negative partial sums there corresponds (by a finite reordering) a unique sequence with first (absolute) minimum at the \( k \)th place. The partial sums \( s_0 = 0, s_1, s_2, \ldots \) of the rearranged sequence consist of a first part \( s_0, s_1, \ldots, s_k \) and a second part \( s_{k+1}, s_{k+2}, \ldots \) such that the partials sums satisfy \( s_i > s_k \) for \( i \leq k \) and \( s_i - s_k \geq 0 \) for \( i > k \). For a random walk the joint distribution of the \( X_i \) is invariant under finite permutations, and the two parts are independent. The first part has probability

\[
P(0 > S_k, S_1 > S_k, \ldots, S_{k-1} > S_k) = P(S_1 < 0, \ldots, S_k < 0)
\]

(by reversing the order of the variables), the second part has probability

\[
P(S_{k+1} - S_k \geq 0, S_{k+2} - S_k \geq 0, \ldots) = P(S_1 \geq 0, S_2 \geq 0, \ldots) = 1 - r(1).
\]

This yields the first equation of (4.1). The second one follows immediately from the factorization identity \((1 - a(z))(1 - r(z)) = 1 - z\) (recall Theorem 4.2) and the third one from Sparre Andersen’s theorem.

In some cases \( r(z) \) can be computed in closed form, and the asymptotics of \( Z_0 \) can be obtained from an explicit formula. An example is the normal random walk. Let the iid steps \( X_i \) be \( N(\delta, \sigma^2) \)-distributed. Here we only assume that \( \delta \neq 0 \), i.e., we consider the cases of positive and negative \( \delta \) simultaneously and let \( d := |\delta|, q := \frac{\delta^2}{2\sigma^2} \).

Example 4.5. For the normal random walk we have

(a) \( r(z) = 1 - (1 - z)^{1/2} \exp \left( \frac{\text{sign}(\delta) d^2}{\pi \sigma^2} \int_0^1 \int_0^\infty e^{-d^2(y^2 + x^2)/2\sigma^2} \frac{1}{1 - z e^{-d^2(y^2 + x^2)/2\sigma^2}} dy dx \right) \).

(b) \( qZ_0 \longrightarrow A \) in distribution as \( \delta^2/\sigma^2 \searrow 0, \delta \searrow 0 \).

(c) \( r(e^{-\theta q})^{1/\sqrt{q}} \longrightarrow e^{-(\sqrt{1 + \delta} - 1)} \) as \( q \searrow 0, \delta \nearrow 0 \).

Note that here \( \sigma^2 \) may vary with \( \delta \), it is only essential that \( \delta/\sigma \longrightarrow 0 \).

Proof. Directly from Sparre Andersen’s theorem we find that

\[
\log \left( \frac{1}{1 - r(z)} \right) = \sum_{n=1}^{\infty} \frac{z^n}{n} P(S_n < 0) = \sum_{n=1}^{\infty} \frac{z^n}{n} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2n\pi \sigma^2}} e^{-x^2/2n\sigma^2} dx
\]

\[
= \sum_{n=1}^{\infty} \frac{z^n}{n} \left( \frac{1}{2} - \text{sign}(\delta) \int_0^{\infty} \frac{1}{\sqrt{2n\pi \sigma^2}} e^{-x^2/2n\sigma^2} dx \right).
\]

Hence,

\[
1 - r(z) = (1 - z)^{1/2} \exp(\text{sign}(\delta) G(z)),
\]

(4.2)

where

\[
G(z) = \sum_{n=1}^{\infty} \frac{z^n}{n} \int_0^{\infty} \frac{1}{\sqrt{2n\pi \sigma^2}} e^{-x^2/2n\sigma^2} dx.
\]
We have
\[
\int_0^{nd} \frac{1}{\sqrt{2n\pi\sigma^2}} e^{-x^2/2n\sigma^2} \, dx = \int_0^n \frac{1}{\sqrt{2\pi\sigma^2}} e^{-ny^2/2\sigma^2} \, dy = \frac{n}{\pi\sigma^2} \int_0^{\infty} e^{-n(y^2+x^2)/2\sigma^2} \, dy \, dx
\]
and therefore
\[
G(z) = \frac{d^2}{\pi\sigma^2} \int_0^{\infty} \int_0^{\infty} e^{-d^2(y^2+x^2)/2\sigma^2} \, dy \, dx \cdot \left(1 - ze^{-d^2(y^2+x^2)/2\sigma^2}\right) \, dx,
\]
proving (a). Note that $G(z)$ depends only on the ratio $q = d^2/2\sigma^2$. Fix $s > 0$. Setting $z = e^{-qs}$ we obtain for $q \downarrow 0$ (by dominated convergence):
\[
G(e^{-qs}) = 2\pi \int_0^{\infty} \int_0^{\infty} e^{-d^2(y^2+x^2)/2\sigma^2} \, dy \, dx \cdot \left(1 + \sqrt{1 + s\sqrt{s}}\right).
\]
From this (b) and (c) follow easily.

It is of methodological interest to have also a purely fluctuation-theoretic proof of Theorem 3.4, i.e., a proof which does not rely on the “functional limit theorem” approach used above. The reviewer suggested the following alternative derivation of 3.4 based on Theorem 4.4. Assume the conditions introduced in Section 3.

**Theorem 4.6. = Theorem 3.4**
\[
\frac{\delta^2}{2\sigma^2(\delta)} Z_0^\delta \overset{\text{dist}}{\longrightarrow} A \text{ in distribution as } \delta \downarrow 0.
\]

**Proof.** In principle, we follow the line of argument used for a similar proof in [20]. Let $\varepsilon > 0$ and split the series in the exponent of the right-hand side of (4.1) into three parts:
\[
\sum_{k=1}^{\infty} \varepsilon/\delta^2 + \sum_{T/\delta^2} + \sum_{T/\delta^2} = \sum_1 + \sum_2 + \sum_3.
\]
Let $s > 0$ and set
\[
z = e^{-s\delta^2/2\sigma^2(\delta)}.
\]
We consider the different sums separately, starting with $\sum_1$:
\[
\sum_{k=0}^{\varepsilon/\delta^2} (1-z^k) \frac{\mathbb{P}(S_k^\delta < 0)}{k} \leq \frac{s\delta^2}{2\sigma^2(\delta)} \sum_{k=0}^{\varepsilon/\delta^2} \mathbb{P}(S_k^\delta < 0) \leq \frac{s\varepsilon}{2\sigma^2(\delta)}.
\]
Furthermore, $\mathbb{P}(S_k^\delta < 0) = \mathbb{P}(\sum_{j=1}^{k} X_{\delta,j} < -k\delta) \leq \sigma^2(\delta)/(k\delta^2)$ by Chebyshev’s inequality. Therefore we obtain for $\varepsilon > \delta^2$
\[
\sum_{k \geq \varepsilon/\delta^2} \frac{1-z^k}{k} \mathbb{P}(S_k^\delta < 0) \leq \frac{\sigma^2(\delta)}{\delta^2} \sum_{k \geq \varepsilon/\delta^2} \frac{1}{k^2} \leq \frac{\sigma^2(\delta)}{\delta^2} \int_{\varepsilon/\delta^2}^{\infty} \frac{1}{x-1} \, dx = \frac{\sigma^2(\delta)}{\varepsilon - \delta^2}.
\]
Since $\sigma(\delta) \to \sigma^2 \in (0, \infty)$ as $\delta \to 0$, there is a $\delta_0$ such that $2\delta_0^2 < \varepsilon$ and $\sigma^2(\delta)$ is bounded for $\delta \leq \delta_0$. Without loss of generality assume in the sequel $\delta \leq \delta_0$. Then $\sum_3$ can be made arbitrarily small by a suitable choice of $T$, and $\sum_2 \leq 2C/\varepsilon$ for a suitable constant $C$.

For $\sum_2$ we use the asymptotic normality of $\delta S_{t/\delta^2}$ (which is implied by the Lindeberg condition, see the beginning of section 3):

$$\mathbb{P}(\delta S_{k}^d < 0) \to \mathbb{P}(N(t, \sigma^2 t) < 0) = \Phi(-\frac{t}{\sigma^2}) as \delta \to 0, k\delta^2 \to t$$

(uniformly for $t \in [\varepsilon, T]$), and by the dominated convergence we conclude that

$$\sum_2 \to \int_{\varepsilon}^{T} \frac{1 - e^{-t/2\sigma^2}}{t} \Phi(-\sqrt{\frac{t}{\sigma^2}}) dt.$$ 

Letting $\varepsilon \to 0, T \to \infty$ we finally arrive at

$$\mathbb{E}e^{-\frac{s^2}{2\sigma^2}} \to \exp\left\{ -\int_{0}^{\infty} \frac{1}{u} \Phi(-\sqrt{\frac{tu}{2}}) du \right\}.$$ 

(4.3)

Evaluating the integral finishes the proof. Avoiding the calculation, it suffices to notice that the right side of (4.3) is independent of the underlying distribution of the random walk so that one can look at the example of the normal random walk computed above, which leads to the conclusion that the right side of (4.3) is equal to $2/(1 + \sqrt{1 + s})$.

The advantage of this proof is that it generalizes to the $\alpha$-stable case ($1 < \alpha < 2$) essentially unchanged - the main difficulties (the corresponding estimates for these cases) can be overcome using inequality (6) in [20].

We close this section with a few remarks on the simple random walk taking step $+1$ with probability $p > 1/2$ and step $-1$ with probability $q = 1 - p$. It is well-known that in this example

$$r(z) = \frac{1 - \sqrt{1 - 4pqz^2}}{2pz},$$

so that a quick calculation shows that

$$\mathbb{E}Z_0 = \frac{1 - r(1)}{1 - r(z)} = \frac{(p - q)(1 + \sqrt{1 - 4pqz^2})}{p(1 - 2z^2 + \sqrt{1 - 4pqz^2})}$$

and $2(p - \frac{1}{2})^2Z_0 \to A$ in distribution as $p \searrow 1/2$.

**Remark 4.7.** Let $T_0(p) = \sup\{n \geq 0 : S_n^{(0)} = 0\}$ the time of the last return to the origin. In the symmetric case $p = 1/2$ the walk is persistent and $T_0(1/2) = \infty$ almost surely. In the transient case $p > 1/2$, $T_0(p)$ has generating function

$$h(z) = \frac{p - q}{\sqrt{1 - 4pqz^2}}.$$ 

A short computation yields that $\frac{1}{2}(p - q)^2T_0(p)$ converges in distribution as $p \searrow 1/2$, the limiting distribution having Laplace transform $\frac{1}{\sqrt{1 + s}}$, i.e., being the $\Gamma_{1, \frac{1}{2}}^1$ distribution with density $\gamma_{1, \frac{1}{2}}(t)$ as above.
Remark 4.8. Let $N_0(p)$ denote the number of zeros of the random walk. Then

$$P(N_0(p) = r, T_0(p) = 2n) = \frac{r}{n-r} \binom{2n-r}{n} 2^{pq}$$

and $(\delta N_0(p), \frac{1}{2} \delta^2 T_0(p))$ converges weakly to the distribution with density

$$f(y,t) = 1_{(0,\infty)}(y) 1_{(0,\infty)}(t) \frac{1}{2t} \frac{1}{\sqrt{2\pi} t} e^{-\frac{y^2}{4t}} e^{-t}.$$

In particular, $\delta N_0(p)$ is asymptotically $\exp(1)$. For the symmetric random walk let $N_{0,2n}$ denote the number of zeros up to time $2n$. A classical theorem of Chung-Hunt [5] states that $\sqrt{2/n} N_{0,2n}$ is asymptotically distributed as $|N(0,1)|$. All these results show that deviations from the symmetric random walk become clearly visible after $n \approx \delta^{-2}$ steps. While characteristics like the positive sojourn time and the last exit time from zero are in both cases of approximately the same size their distributions differ. For the last exit time from zero a precise description is given in Theorem 2.3.

 Apparently the distribution of $A$ occurs naturally as a limit of occupation times for random walks with drift. It is well-known (see e.g. Section XIV.3 in [8]) that the deeper reason for the frequent occurrence of the (generalized) arcsine distributions lies in their intimate connection to distribution functions with regularly varying tails. The same explanation applies here. In the case of drift zero the distribution functions of the ladder epochs are attracted to the standard positive stable distribution of index $1/2$ and the positive (negative) sojourn times are asymptotically arcsine-distributed. In the cases with small drift (and finite variance) the ladder epochs are attracted to an associated distribution of this stable distribution, and therefore the positive (negative) sojourn times have asymptotically the distribution of $A$.

Acknowledgement. We would like to thank the referee for valuable remarks and in particular for suggesting the alternative proof of Theorem 3.4 given in Section 4.

References

[1] P. Billingsley (1999) Convergence of Probability Measures, Second Edition. John Wiley&Sons.
[2] A.N. Borodin and P. Salminen (2002) Handbook of Brownian Motion - Facts and Formulae, Second Edition. Birkhäuser.
[3] O.J. Boxma and J.W. Cohen (1999) Heavy-traffic analysis for the GI/G/1 queue with heavy-tailed distributions. Queueing Syst. 33, 177-204.
[4] M. Brown, E.A. Peköz and S.M. Ross (2010) Some results for skipfree random walk. Prob. Engineer. Inform. Sci. 24, 491-507.
[5] K.L. Chung and G.A. Hunt (1949) On the zeros of $\sum (\pm 1)^n$. Ann. Math. 50, 385-400.
[6] R.A. Doney and M. Yor (1998) On a formula of Takács for Brownian motion with drift. J. Appl. Prob. 35, 272-280.
[7] W. Feller (1968) An Introduction to Probability Theory and Its Applications, Vol. I, Third Edition. John Wiley&Sons.
[8] W. Feller (1971) An Introduction to Probability Theory and Its Applications, Vol. II, Second Edition. John Wiley&Sons.
[9] J. Hoffmann-Jorgensen (1999) The arcsine law. J. Theoret. Prob. 12, 131-145.
[10] J.-P. Imhof (1986) On the time spent above a level by Brownian motion with negative drift. Adv. Appl. Prob. 18, 1017-1018.
[11] J.-P. Imhof (1999) On some equalities of laws for Brownian motion with drift. J. Appl. Prob. 36, 682-687.
[12] S. Karlin and H. M. Taylor (1975) A first Course in Stochastic Processes, Second Edition. Academic Press.
[13] J.F.C. Kingman (1965) The heavy traffic approximation in the theory of queues. In: W.L. Smith and W.E. Wilkinson (eds), Proceedings of Symposium on Congestion Theory, 137-159. University of North Carolina Press.

[14] K. M. Kosinski, O.J. Boxma and A.P. Zwart (2011) Convergence of the all-time supremum of a Levy process in the heavy-traffic regime. Queueing Systems 67, 295-304.

[15] A. Lachal (2012) Sojourn time in $\mathbb{Z}^+$ for the Bernoulli random walk on $\mathbb{Z}$. ESAIM: P & S 16, 324-351.

[16] P. Marchal (1998) Temps d’occupation de $(0, \infty)$ pour les marches aléatoires. Stochastics Stochastics Reports 64, 267-282.

[17] Yu. V. Prohorov (1963) Transition phenomena in queueing processes. I. (in Russian). Litovsk. Mat. Sb 3, 199-205.

[18] S. Resnick and G. Samorodnitsky (2010) A heavy traffic limit theorem for workload processes with heavy tailed service requirements. Management Sci. 46, 1236-1248.

[19] P. Salminen (1988) On the first hitting time and the last exit time of Brownian motion to/from a moving boundary, Adv. Appl. Prob. 20, 411-426.

[20] S. Shneer and V. Wachtel (2010) Heavy-traffic analysis of the maximum of an asymptotically stable random walk. Theory Probab. Appl. 55, 332-341.

[21] L. Takács (1996) On a generalization of the arc-sine law. Ann. Appl. Prob. 6, 1035-1040.

Bundesamt für Sicherheit in der Informationstechnik (BSI), Godesberger Allee 185–189, 53175 Bonn, Germany

E-mail address: ernst.schulte-geers@bsi.bund.de

Institute of Mathematics, University of Osnabrück, 49069 Osnabrück, Germany

E-mail address: wstadje@uos.de