A note on the simple random walk on \( \mathbb{Z}^2 \): probability of exiting sequences of sets

Stanislav Volkov\(^*\)

December 20, 2006

Abstract

In this note we establish that the probability that the simple random walk on \( \mathbb{Z}^2 \) returns to its origin before leaving a strip of width \( L \) has asymptotically the same probability as the one for hitting the origin before exiting the centered box of the same size. We also generalize this theorem for fairly arbitrary sequences of increasing sets in \( \mathbb{Z}^2 \).

Mathematics Subject Classification: primary 60J10, secondary 60D05.

Keywords: simple random walk, hitting probabilities.

1 Introduction and main results

Let \( (X_n, n \in \mathbb{Z}^+; \mathbb{P}_x, x \in \mathbb{Z}^2) \) be a simple random walk on \( \mathbb{Z}^2 \) started at \( X_0 = x \). Throughout the paper, for \( x \in \mathbb{Z}^2 \) the pair \((x_1, x_2)\) denotes the coordinates of \( x \). For \( a, b > 0 \) let

\[
R_{a,b} = \{ x \in \mathbb{Z}^2 : |x_1| \leq a, |x_2| \leq b \}
\]

be a rectangle centered at the origin \( 0 = (0,0) \) and

\[
S_b = \{ x \in \mathbb{Z}^2 : |x_2| \leq b \} = R_{\infty,b}
\]

\(^*\)Department of Mathematics, University of Bristol, BS8 1TW, U.K.
E-mail: S.Volkov@bristol.ac.uk
be a strip of width $2b$.

Let $x \in V \subset \mathbb{Z}^2$. Following van den Berg (2004), we want to study the probability of exiting $V$ before returning to $x$, i.e.

$$q_x[V] = \mathbb{P}_x \left( \tau(\mathbb{Z}^2 \setminus V) < \tau(x) \right)$$

where for any $U \subset \mathbb{R}^2$ we define the stopping time as

$$\tau(U) = \inf\{n \geq 1 : X_n \in U\}.$$  

This problem has links with other important problems: see van den Berg (2004) and references therein.

It was shown in van den Berg (2004) that in the case when $V$ is a square or a slightly elongated rectangle centered around the origin,

$$\lim_{L \to \infty} q_0[R_{L,L}] \log(L) = \frac{\pi}{2}. \tag{1}$$

Our purpose is to extend this result to a rectangle of any shape as well as to an infinite strip. The next statement tells us that this probability essentially depends only on how fast the smaller side of the rectangle grows, thus strengthening the results of Example 3 from van den Berg (2004).

**Theorem 1** (a) Let $i_0$ be a positive integer. Let $a_i$ and $b_i$, $i = 1, 2, \ldots$ be two sequences of positive integers, such that

$$\lim_{i \to \infty} b_i = \infty, \quad a_i \geq b_i \text{ for all } i \geq i_0.$$  

Then

$$\lim_{i \to \infty} q_0[R_{a_i,b_i}] \log(b_i) = \frac{\pi}{2}.$$  

(b)

$$\lim_{L \to \infty} q_0[S_L] \log(L) = \frac{\pi}{2}.$$  

Note that by (1) part (a) of the Theorem will automatically follow from part (b), since

$$q_0[S_L] \leq q_0[R_{a,L}] \leq q_0[R_{L,L}]$$
as long as \( a \geq L \).

The following statement is based on Lemma 22.1 in Révész (1990).

**Lemma 1** Let

\[
C(r) = \{ x \in \mathbb{Z}^2 : \| x \| \leq r \},
\]

where \( \| \cdot \| \) denotes the usual Euclidean norm, be “a ball” of radius \( r \) in \( \mathbb{Z}^2 \). Then

\[
\lim_{r \to \infty} q_0[C(r)] \log(r) = \frac{\pi}{2}.
\]

Following the lines of Lemma 5 in van den Berg (2004), we will prove

**Lemma 2** Let \( \alpha \) and \( \beta \) be two positive constants. Then

\[
\lim_{L \to \infty} q_0[R_{\alpha L, \beta L}] \log(L) = \frac{\pi}{2}.
\]

**Proof.** Without loss of generality suppose \( \alpha \geq \beta \). Exactly as in van den Berg (2004), observe that

\[
C(\beta L) \subset R_{\alpha L, \beta L} \subset C(\sqrt{\alpha^2 + \beta^2}L)
\]

whence

\[
q_0[C(\beta L)] \geq q_0[R_{\alpha L, \beta L}] \geq q_0[C(\sqrt{\alpha^2 + \beta^2}L)]
\]

Now we multiply this by \( \log(L) \), take into account that \( \log(\alpha L)/\log(L) = 1 + o(1) \) and \( \log(\sqrt{\alpha^2 + \beta^2}L)/\log(L) = 1 + o(1) \), let \( L \to \infty \), and finally use Lemma 1.

The following proof uses combinatorial arguments, and though it may be obtained from a stronger Theorem 2, we present it for the expository purposes.

**Proof of Theorem 1.** As we mentioned before, it is sufficient to prove just part (b). Fix \( \alpha > 2, \beta \equiv 1, \) and a positive \( \varepsilon < 1 \). Then by Lemma 2, there is \( L_0 = L_0(\alpha, 1, \varepsilon) \) such that for all \( L \geq L_0 \)

\[
q_0[R_{\alpha L, L}] \geq \frac{\pi}{2 \log(L)}(1 - \varepsilon).
\]
Let
\[ E_s = \{ \tau (\mathbb{Z}^2 \setminus S_L) < \tau(0) \} \]
be the event of exiting the strip before hitting 0, and \[ E_h = \{ \tau (\mathbb{Z}^2 \setminus R_{\alpha L, L}) < \tau(0) \} \]
\[ \cap \{ \tau (x \in \mathbb{Z}^2 : |x_1| = |\alpha L + 1|) \geq \tau (x \in \mathbb{Z}^2 : |x_2| = L + 1) \}, \]
\[ E_v = \{ \tau (\mathbb{Z}^2 \setminus R_{\alpha L, L}) < \tau(0) \} \]
\[ \cap \{ \tau (x \in \mathbb{Z}^2 : |x_1| = |\alpha L + 1|) < \tau (x \in \mathbb{Z}^2 : |x_2| = L + 1) \} \]
(3)
be the events of exiting \( R_{\alpha L, L} \) before hitting 0 in such a way that the horizontal (vertical resp.) side of this rectangle is crossed first. Here \( \lfloor \cdot \rfloor \) denotes the integer part of its argument. Then
\[ P_0(E_s) = P_0(E_s \cap E_h) + P_0(E_s \cap E_v) + P_0(E_s \cap (E_v \cup E_h)^c) \]
\[ = P_0(E_h) + P_0(E_s | E_v)P_0(E_v) = P_0(E_h \cup E_v) - P_0(E_s | E_v)P_0(E_v) \]
\[ \geq q_0 [R_{\alpha L, L}](1 - P_0(E_s | E_v)) \]
(4)
since \( E_h \subset E_s \), \( P_0(E_h \cup E_v) = q_0 [R_{\alpha L, L}] \) and on the event \((E_v \cup E_h)^c\) point 0 is hit before exiting \( R_{\alpha L, L} \subset S_L \). Next we will show that \( P_0(E_s^c | E_v) \) is close to 0. Indeed, on \( E_v \), the walk must exit through the left or right side of the rectangle \( R_{\alpha L, L} \). Let
\[ y^{(1)} := X_{\tau(\mathbb{Z}^2 \setminus R_{\alpha L, L})} \in \{ x : x_1 = \pm |\alpha L + 1|, |x_2| \leq L \} \]
be the point of this exit. Now consider the simple random walk started at \( y^{(1)} \) when it exits the \((2L) \times (2L)\) square
\[ y^{(1)} + R_{L,L} \equiv \{ y \in \mathbb{Z}^2 : (y_1 - y^{(1)}_1, y_2 - y^{(1)}_2) \in R_{L,L} \} \]
centered at \( y^{(1)} \). Note that either the upper or lower sides of this square must not belong to \( S_L \). Let \( E_1 \) be the event that it exits this square via
not belonging to $S_L$ side. By symmetry, $\mathbb{P}_{g(1)}(E_1) = 1/4$. Let $y^{(2)}$ be the point where the walk exits the above square. Now consider another square $y^{(2)} + R_{L,L}$ centered at $y^{(2)}$, and so on – that is, recursively define $E_k$’s as the events that the walk exits the square $y_k + R_{L,L}$ via the side which does not intersect with $S_L$ (for definiteness, if there are more than one, we set it to be the upper side only), and in any case call the point of exit $y^{(k+1)}$.

Observe that the events $E_1, E_2, \ldots, E_k$ are independent and all have probability 1/4. At the same time, if an event $E_k$ occurs for some $k \leq \alpha - 1$, this implies that 0 was definitely not yet hit and thus $E_s$ occurs. Therefore,

$$
\mathbb{P}_0(E_s^c | E_v) \leq \mathbb{P}_0 \left( \bigcap_{k=1}^{\lfloor \alpha - 1 \rfloor} E_k^c \right) = \prod_{k=1}^{\lfloor \alpha - 1 \rfloor} \mathbb{P}_{g^{(k)}}(E_k^c) \leq \left( \frac{3}{4} \right)^{\alpha - 2}
$$

Consequently, from equations (2) and (4) it follows that

$$
q_0 [S_L] \geq \pi \frac{2}{\log(L)} \left( 1 - \varepsilon \right) \left( 1 - \left( \frac{3}{4} \right)^{\alpha - 2} \right).
$$

By choosing $\alpha$ large and $\varepsilon$ small we establish that $q_0$ is

$$
\lim \inf_{L \to \infty} q_0 [S_L] \log(L) \geq \frac{\pi}{2}.
$$

On the other hand, by (1)

$$
\lim \sup_{L \to \infty} q_0 [S_L] \log(L) \leq \lim \sup_{L \to \infty} q_0 [R_{L,L}] \log(L) = \frac{\pi}{2}.
$$

which finishes the proof. 

2 Generalizations

For any set $A \subset \mathbb{Z}^2$ let

$$
\partial A = \{ y \in \mathbb{Z}^2 \setminus A : \| y - x \| = 1 \text{ for some } x \in A \}
$$

be the discrete border of this set. Now, recall that $C(r)$ is a circle of radius $r$, and let $\partial C(r)$ be its discrete border. Clearly, the number of points in $\partial C(r)$ is of order $r$. The next statement is applicable to a sequence of sets of arbitrary shape.
**Theorem 2** Let $V_i$ be a sequence of subsets of $\mathbb{Z}^2$ such that there are two positive sequences $\{a_i\}, \{b_i\}$ with the following properties:

\[
\lim_{i \to \infty} a_i = \infty,
\]

\[
\lim_{i \to \infty} \frac{\log(a_i)}{\log(b_i)} = 1,
\]

\[
C(a_i) \subseteq V_i,
\]

\[
\lim_{i \to \infty} \frac{|\partial C(b_i) \cap V_i|}{b_i} = 0.
\]

Then

\[
\lim_{i \to \infty} q_0[V_i] \log(a_i) = \frac{\pi}{2}.
\]

Before we proceed with the proof, we restate Lemma 1.7.4 from Lawler (1991).

**Lemma 3** There are two constants $c_1$ and $c_2$, such that for the simple random

\[
c_1/r \leq \mathbb{P}_0(X_{\tau(\partial C(r))} = y) \leq c_2/r
\]

for any $y \in \partial C(r)$.

The next statement is essentially “trivial”.

**Lemma 4** Consider a subset $A \subset \mathbb{Z}^2$ containing 0, and let $\partial A$ be its border. If $\tau^* = \tau(\partial A) = \tau(\mathbb{Z}^2 \setminus A)$ denotes the time of exit from $A$, then

\[
\mathbb{P}_0(X_{\tau^*} \in B \mid \tau^* < \tau(0)) = \mathbb{P}_0(X_{\tau^*} \in B)
\]

for any $B \subseteq \partial A$.

**Proof.** By strong Markov property,

\[
\mathbb{P}_0(X_{\tau^*} \in B \mid \tau^* > \tau(0)) = \mathbb{P}_{X_{\tau(0)}}(X_{\tau^*} \in B) = \mathbb{P}_0(X_{\tau^*} \in B).
\]
Since $\tau^* \neq \tau(0)$, (6) immediately follows.

Proof of Theorem 2. Since $C(a_i) \subseteq V_i$, we have $q_0[C(a_i)] \geq q_0[V_i]$, hence
\[
\limsup_{i \to \infty} q_0[V_i] \log(a_i) \leq \limsup_{i \to \infty} q_0[C(a_i)] \log(a_i) = \frac{\pi}{2}
\]
by Lemma 1.

Next, let
\[
H_i = \{X_{\tau(\partial C(b_i))} < \tau(0)\}
\]
be the event that the simple random walk hits $\partial C(b_i)$ before returning to $0$. Then
\[
q_0[V_i] = \mathbb{P}_0(\tau(\mathbb{Z}^2 \setminus V_i) < \tau(0) \geq \mathbb{P}_0(\tau(\mathbb{Z}^2 \setminus V_i) < \tau(0) | H_i) \mathbb{P}_0(H_i)
\]
\[
\geq \mathbb{P}_0(X_{\tau(\partial C(b_i))} \in \partial C(b_i) \setminus V_i | H_i) \mathbb{P}_0(H_i)
\]
(by Lemma 4)
\[
= \mathbb{P}_0(X_{\tau(\partial C(b_i))} \in \partial C(b_i) \setminus V_i) \mathbb{P}_0(H_i)
\]
\[
= \left[1 - \sum_{y \in C(b_i) \cap V_i} \mathbb{P}_0(X_{\tau(\partial C(b_i))} = y)\right] \mathbb{P}_0(H_i)
\]
(by Lemma 3)
\[
\geq \left(1 - \frac{c_1 |\partial C(b_i) \cap V_i|}{b_i} \right) q_0[C(b_i)],
\]
since also $\mathbb{P}_0(H_i) = q_0[C(b_i)]$. Now applying Lemma 1, we conclude that
\[
\liminf_{i \to \infty} q_0[V_i] \log(a_i) = \liminf_{i \to \infty} q_0[V_i] \log(b_i)
\]
\[
\geq \liminf_{i \to \infty} \left(1 - \frac{c_1 |\partial C(b_i) \cap V_i|}{b_i} \right) q_0[C(b_i)] \log(b_i)
\]
\[
= \frac{\pi}{2}.
\]
Note that for the sequence of strips $S_i$ the conditions of Theorem 2 are fulfilled if we choose $a_i = i$, $b_i = i \log(i)$, since $|C(b_i) \cap V_i| \leq 4i+2$. Therefore, one can obtain Theorem 1 as a corollary of a more general statement.

Another example where one can apply Theorem 2, is the following problem. Let

$$V = \{ x \in \mathbb{Z}^2 : x_2 \geq x_1^2 \}$$

be the interior of a parabola, the walk starts at $x^{(i)} = (0, i)$, $i > 0$, and we are interested in the asymptotical probability of exiting $V$ before hitting the vertex where the walk has originated, that is $q_{x^{(i)}}[V]$. To solve it, observe that $q_{x^{(i)}}[V] = q_0[V_i]$ where

$$V_i = \{ x \in \mathbb{Z}^2 : x_2 \geq x_1^2 - i \}$$

is the parabola shifted down by $i$. Now Theorem 2 applies with

$$a_i = \sqrt{i - \frac{1}{4}}, \quad b_i = \sqrt{i \log(i)} + i$$

since an easy calculation shows that $|C(b_i) \cap V_i| \leq 4\sqrt{i}$.

### 2.1 Exiting other sets

Not for all sequences $V_L$ of subsets of $\mathbb{Z}^2$ one can apply Theorem 2 directly, yet still the same result about the asymptotical behaviour of the hitting probabilities might still hold. An interesting example is the half-plane

$$H = \{ x \in \mathbb{Z}^2 : x_2 \geq 0 \},$$

where we start the walk at the points $(0, L)$, $L \in \mathbb{Z}_+$. First, we shift the half plane down such that the walk for any $L$ starts at $0$, and will rather study $q_0[H_L] = q_{(0,L)}[H]$ where

$$H_L = \{ x \in \mathbb{Z}^2 : x_2 \geq -L \}$$
Clearly, the conditions of Theorem 2 cannot be satisfied, as for any increasing sequence of \(b_i\)'s the limit in (5) will be \(1/2\) and not 0 as required. Still, the following is true.

**Proposition 1**

\[
\lim_{L \to \infty} q_{0(L)}[H \log(L)] = \lim_{L \to \infty} q_{0(L)}[H_L \log(L)] = \frac{\pi}{2}.
\]

**Proof.** Let

\[M = M_L = \lfloor L \log(L) \rfloor\]

and consider a circular segment which is a slice of a circle of radius \(M\) centered at \((0, -2L)\):

\[C^* = C^*_L = \{x : x_1^2 + (x_2 + 2L)^2 \leq M^2, \ x_2 \geq -L\}.
\]

Then Theorem 2 applies to the sequence of \(C^*_i\)'s with \(a_i = i, b_i = 2i + M_i\), hence

\[
\lim_{L \to \infty} q_{0(C^*_L)}[\log(L)] = \frac{\pi}{2}.
\]  

(7)

Let \(\tau^* = \tau_{\infty}^* = \tau(\mathbb{Z}^2 \setminus C^*_L)\) be the time of the exit from \(C^*\), and also let

\[A = A_L = \{x \in \partial C^*_L : x_2 = -L - 1\}\]

be the border of the bottom flat side of \(C^*\).

Denote the cartesian coordinates of the walk as \(X_n = ([X_n]_1, [X_n]_2)\). Let

\[
\xi_n = \begin{cases} 
\log(((X_n)_1 - a)^2 + ((X_n)_2 - b)^2 - \frac{1}{2}), & \text{if } X_n \neq (a, b), \\
-\infty, & \text{if } X_n = (a, b).
\end{cases}
\]

Let \(\mathcal{F}_n = \sigma(X_1, \ldots, X_n)\) be a sigma-algebra generated by the first \(n\) steps of the walk. The proof of the following statement is given after the proof of the Proposition.

**Claim 1** Suppose \(X_n \notin \{(a, b), (a + 1, b), (a - 1, b), (a, b + 1), (a, b - 1)\}\). Then

\[
E(\xi_{n+1} - \xi_n | \mathcal{F}_n) \leq 0.
\]
Now set \( a = 0 \) and \( b = -2L \). Since neither \((0, -2L)\) nor \((-1, -2L)\) nor \((0, -2L \pm 1)\) belong to \( C^* \), by Claim 1 we conclude that \( \xi_{n \wedge \tau^*} \) is a supermartingale with respect to filtration \( \mathcal{F}_n \), and then by a corollary of the optional stopping theorem (see Durrett (1996), p. 273) and letting \( n \to \infty \),

\[
\mathbb{E} \xi_{\tau^*} \leq \xi_0 = \log(4L^2) + o(1).
\]

We will use this formula to estimate the probability that the walk exits \( C^* \) via the bottom flat side of \( C^* \). We split the probability space into two events: \( X_{\tau^*} \in A_L \) and \( X_{\tau^*} \notin A_L \) and recompute \( \mathbb{E} \xi_{\tau^*} \):

\[
\mathbb{E} \xi_{\tau^*} = \mathbb{E} (\xi_{\tau^*} \mid [X_{\tau^*}]_2 = -L - 1) \mathbb{P}(X_{\tau^*} \in A_L)
+ \mathbb{E} (\xi_{\tau^*} \mid [X_{\tau^*}]_2 \geq -L)(1 - \mathbb{P}(X_{\tau^*} \in A_L))
\geq (\log(L^2) + o(1)) \mathbb{P}(X_{\tau^*} \in A_L) + (\log(M^2) + o(1)) \left(1 - \mathbb{P}(X_{\tau^*} \in A_L)\right),
\]

since when the exit occurs via the arc, \( \xi_{\tau^*} = \log M^2 + O(1/M^2) \), and when the exit occurs via the chord, \( \xi_{\tau^*} \geq \log[-2L - (-L - 1)]^2 \). Therefore,

\[
\mathbb{P}(X_{\tau^*} \in A_L) \geq \frac{\log(M^2/(4L^2))}{\log(M^2/L^2)} + o(1) = 1 - \frac{2}{\log \log L} + o(1) \to 1 \quad \text{as} \quad L \to \infty.
\]

Next, since \( \mathbb{P}(\tau^*_L < \tau(0)) = q_0[C^*_L] \),

\[
q_0[H_L] \geq \mathbb{P}(\tau(Z^2 \setminus H_L) < \tau(0) \text{ and } \tau^*_L < \tau(0))
= \mathbb{P}(\tau(Z^2 \setminus H_L) < \tau(0) \mid \tau^*_L < \tau(0)) q_0[C^*_L]
\geq \mathbb{P}(X_{\tau^*_L} \in A_L \mid \tau^*_L < \tau(0)) q_0[C^*_L]
\geq \mathbb{P}(X_{\tau^*_L} \in A_L) q_0[C^*_L] = (1 - o(1)) \times q_0[C^*_L].
\]

Combining this with (7) and an obviously inequality \( q_0[H_L] \leq q_0[C^*_L] \) yields the statement of Proposition. \( \blacksquare \)
Proof of Claim 1. First of all, observe that we can set \(a=b=0\) without loss of generality. Suppose now that \(X_n = (x, y) \in \mathbb{Z}^2\), where \(x^2 + y^2 > 1\), so that \(\xi_n = \log(x^2 + y^2 - 1/2)\), and compute \(\mathbb{E}(\xi_{n+1} - \xi_n \mid \mathcal{F}_n)\) as follows:

\[
\mathbb{E}(\xi_{n+1} - \xi_n \mid \mathcal{F}_n) = \frac{1}{4} \left[ \log((x+1)^2 + y^2 - 1/2) + \log((x-1)^2 + y^2 - 1/2) \right. \\
+ \left. \log(x^2 + (y+1)^2 - 1/2) + \log(x^2 + (y-1)^2 - 1/2) \right] \\
- \log(x^2 + y^2 - 1/2) = \frac{1}{4} \log Q_{x,y}
\]

where

\[
Q_{x,y} = \frac{(2x^2 + 4x + 1 + 2y^2)(2x^2 - 4x + 1 + 2y^2)(2y^2 + 4y + 1 + 2x^2)}{(2y^2 - 4y + 1 + 2x^2)(2x^2 + 2y^2 - 1)^4}
\]

Consequently, \(\frac{1}{4} \log Q_{x,y} \leq 0\) whenever \(\log Q_{x,y}\) is defined (iff \(Q_{x,y} > 0\)). It is easy to check now that \(\log Q_{x,y}\) is indeed defined unless \((x, y) \in \{(0, 0), (1, 0), (0, 1), (-1, 0), (0, -1)\}\).

Finally, we present an open problem: what is the probability of hitting the half line before returning to the origin? Namely, suppose that 

\[ V_L = \mathbb{Z}^2 \setminus \{ x : x_1 < -L, x_2 = 0 \}. \]

What is the asymptotical behaviour of \(q_0[V_L]\)?

Note that it is clear that \(\limsup q_0[V_L] \log(L) \leq \pi/2\) since \(C(L) \subset V_L\). Also, \(\liminf q_0[V_L] \log(L) \geq \pi/4\), since by Proposition 1 the probability that the walk hits the vertical line \(\{ x : x_1 = -L \}\) before returning to the origin, is approximately \(\frac{\pi}{2 \log(L)}\), and then by symmetry, the probability to hit \((-2L, 0) \in \mathbb{Z}^2 \setminus V_L\) before \(0\) is exactly \(1/2\). Thus, it is natural to guess that

\[ \lim_{L \to \infty} q_0[V_L] \log(L) = \rho \frac{\pi}{2} \]

with \(\rho \in [1/2, 1]\); however we do not have a proof of this fact. From discussions with Ofer Zeitouni we conjecture though that \(\rho = 1\) nevertheless.
Acknowledgment

The author would like to thank Michiel van den Berg for the introduction and discussions of this problem, and Greg Lawler for useful suggestions.

References


