Random walks that avoid bounded sets, and applications to the largest gap problem

Vlad Vysotsky

Arizona State University,
St. Petersburg Division of Steklov Institute,
Chebyshev Laboratory at St. Petersburg State University

Darmstadt, July 16, 2014

Outline

I. Random walks that avoid bounded sets

Strongly related to persistence probability. Motivated by Part II.

II. Applications to the largest gap problem

Interesting by itself. Related to persistence probability of iterated random walks.

I. Random walks that avoid bounded sets

1. The exit problem for random walks

Let $X_1, X_2, \ldots$ be i.i.d. r.v.'s so $S_n := x + X_1 + \cdots + X_n$ is a random walk.

Denote $\mathbb{P}_x(\cdot)$ the law of walk starting at x, and put $\mathbb{E}_x f := \int f d\mathbb{P}_x$.

I. Random walks that avoid bounded sets

1. The exit problem for random walks

Let $X_1, X_2, \dots$ be i.i.d. r.v.'s so $S_n := x + X_1 + \cdots + X_n$ is a random walk.
Denote $\mathbb{P}_x(\cdot)$ the law of walk starting at x, and put $\mathbb{E}_x f := \int f d\mathbb{P}_x$.

Let $\tau_B := \inf\{n \geq 1 : S_n \in B\}$ be the hitting time for a Borel set B. A huge number of works is devoted to the asymptotic of $\mathbb{P}_x(\tau_B > n)$ under different assumptions of S_n and B. For example, for $B = (-\infty, 0) \subset \mathbb{R}$ this is the problem of persistence probability. In this case a rather complete theory have been developed (from Sparre-Andersen '50s to Rogozin '72). Some recent advances include exit times from cones in $\mathbb{R}^d$ (Denisov & Wachtel '12+).

We will assume that B is bounded. Fewer results are available here.
Kesten, Spitzer '63: For *any* aperiodic RW in $\mathbb{Z}^{1,2}$ and any *finite* $B \subset \mathbb{Z}^{1,2}$, there exists

$$\lim_{n\to\infty} \frac{\mathbb{P}_x(\tau_B > n)}{\mathbb{P}_0(\tau_{\{0\}} > n)} := g_B(x).$$

Remark: in $\mathbb{Z}^1$, if S_n is centred and asymptotically α-stable with $1 < \alpha \le 2$, then $\mathbb{P}_0(\tau_{\{0\}} > n) \sim cn^{1/\alpha - 1} L(n)$.
Moreover, $L(n) = const$ if $Var(X_1) < \infty$.

We will assume that B is bounded. Fewer results are available here.
Kesten, Spitzer '63: For *any* aperiodic RW in $\mathbb{Z}^{1,2}$ and any *finite* $B \subset \mathbb{Z}^{1,2}$, there exists

$$\lim_{n\to\infty} \frac{\mathbb{P}_x(\tau_B > n)}{\mathbb{P}_0(\tau_{\{0\}} > n)} := g_B(x).$$

Remark: in $\mathbb{Z}^1$, if S_n is centred and asymptotically α-stable with $1 < \alpha \le 2$, then $\mathbb{P}_0(\tau_{\{0\}} > n) \sim cn^{1/\alpha-1}L(n)$.
Moreover, $L(n) = const$ if $Var(X_1) < \infty$.
Remark: $g_B(x)$ is harmonic for the walk killed as it hits B, that is $g_B(x) = \mathbb{E}_x g_B(S_{\tau_B \wedge n})$.

The proof is by induction in $|B|$ and a renewal argument. Neither works in general case.

2. Our assumptions and a lower bound

Assume that the walk is in $\mathbb{R}$, $\mathbb{E}X_1 = 0$, $Var(X_1) := \sigma^2 \in (0, \infty)$. Consider the basic case that $B = (-d, d)$ for some $d > 0$. Put

$$p_n(x) := \mathbb{P}_x(\tau_{(-d,d)} > n), \quad x \notin B.$$

2. Our assumptions and a lower bound

Assume that the walk is in $\mathbb{R}$, $\mathbb{E}X_1 = 0$, $Var(X_1) := \sigma^2 \in (0, \infty)$. Consider the basic case that $B = (-d, d)$ for some $d > 0$. Put

$$p_n(x) := \mathbb{P}_x(\tau_{(-d,d)} > n), \quad x \notin B.$$

Hitting times for half-lines: for *any* $x \geq 0$,

$$\mathbb{P}_x(\tau_{(-\infty,0)} > n) \sim \sqrt{\frac{2}{\pi}} \frac{U_{\geqslant}(x)}{\sigma\sqrt{n}},$$

where $U_{\geqslant}(x)$ is the renewal function. It is harmonic for the walk killed as it enters $(-\infty, 0)$ and satisfies $U_{\geqslant}(x) = \mathbb{E}_x(x - S_{\tau_{(-\infty,0)}})$.

2. Our assumptions and a lower bound

Assume that the walk is in $\mathbb{R}$, $\mathbb{E}X_1 = 0$, $Var(X_1) := \sigma^2 \in (0, \infty)$. Consider the basic case that $B = (-d, d)$ for some $d > 0$. Put

$$p_n(x) := \mathbb{P}_x(\tau_{(-d,d)} > n), \quad x \notin B.$$

Hitting times for half-lines: for *any* $x \geq 0$,

$$\mathbb{P}_x(\tau_{(-\infty,0)} > n) \sim \sqrt{\frac{2}{\pi}} \frac{U_{\geqslant}(x)}{\sigma\sqrt{n}},$$

where $U_{\geqslant}(x)$ is the renewal function. It is harmonic for the walk killed as it enters $(-\infty, 0)$ and satisfies $U_{\geqslant}(x) = \mathbb{E}_x(x - S_{\tau_{(-\infty,0)}})$.
Lower bound: for $|x| \geq d$, staying to one side of B gives

$$p_n(x) \geq \mathbb{P}_x(T_1 > n) \sim \sqrt{\frac{2}{\pi}} \frac{U_d(x)}{\sigma\sqrt{n}}, \quad U_d(x) := \mathbb{E}_x|x - S_{T_1}|,$$

where T_1 is the first moment of jump over $\partial B = \{-d, d\}$.

3. Results for the basic case

Let T_k be the moment of the kth jump over $\{-d, d\}$ *from the outside*; let $H_k := S_{T_k}, k \geq 0$ be the overshoots; denote the # of jumps over $(-d, d)$ before it is hit as $\kappa := \min(k \geq 1 : |H_k| < d)$.

Theorem 1

Let S_n be a random walk with $\mathbb{E}X_1 = 0$, $\mathbb{E}X_1^2 := \sigma^2 \in (0, \infty)$. Then for any $d > 0$ and any $x \in \mathbb{R}$,

$$p_n(x) \sim \sqrt{\frac{2}{\pi}} \frac{V_d(x)}{\sigma\sqrt{n}}, \quad V_d(x) := \mathbb{E}_x\left[\sum_{i=1}^{\kappa} |H_i - H_{i-1}|\right].$$

Moreover, this holds uniformly for $x = o(\sqrt{n})$. Further,

- *$V_d(x)$ is harmonic for the walk killed as it enters $(-d, d)$;*
- *$0 < U_d(x) \leq V_d(x) < \infty$ for $|x| \geq d$;*
- *$V_d(\pm(d + y)) - U_d(\pm(d + y)) \to 0$ as $d \to \infty$ for any $y \geq 0$.*

The later means that there almost no jumps over a wide stripe.

4. Ideas of the proof

1. It costs to jump over:

There exists a $\gamma \in (0,1)$ such that

$$\mathbb{P}_x(|H_1| \geq d) \leq \gamma.$$

This follows since H_1 converge weakly as $x \to \pm\infty$ to the overshoots over "infinitely remote" levels.

4. Ideas of the proof

1. It costs to jump over:

There exists a $\gamma \in (0,1)$ such that

$$\mathbb{P}_x(|H_1| \geq d) \leq \gamma.$$

This follows since H_1 converge weakly as $x \to \pm\infty$ to the overshoots over "infinitely remote" levels.

2. *Regularity of $p_n(x)$* in both x and n is needed.

Lemma: For any $x \in \mathbb{R}$ and $n \geq 1$, $p_n(x) \leq C|x|n^{-1/2}$.

Roughly, $\mathbb{E}_x p_{n-T_1}(H_1)\mathbb{1}_{\{|H_1|\geq d, T_1 \leq n\}}$ is controlled by $\mathbb{E}_x|H_1|$.

4. Ideas of the proof

1. It costs to jump over:

There exists a $\gamma \in (0,1)$ such that

$$\mathbb{P}_x(|H_1| \geq d) \leq \gamma.$$

This follows since H_1 converge weakly as $x \to \pm\infty$ to the overshoots over "infinitely remote" levels.

2. *Regularity of $p_n(x)$* in both x and n is needed.
Lemma: For any $x \in \mathbb{R}$ and $n \geq 1$, $p_n(x) \leq C|x|n^{-1/2}$.
Roughly, $\mathbb{E}_x p_{n-T_1}(H_1)\mathbb{1}_{\{|H_1|\geq d, T_1 \leq n\}}$ is controlled by $\mathbb{E}_x|H_1|$.
3. The mechanism of stabilisation:
For any $\alpha \in (0,1)$ it holds that

$$\mathbb{E}_x|H_1| \leq \alpha|x| + K(\alpha), \quad |x| \geq d.$$

This follows from the known $\mathbb{E}_x|H_1| = o(|x|)$ as $|x| \to \infty$.

5. General sets

Let M be the state space of the random walk, that is $M := \lambda\mathbb{Z}$ if the walk is λ-arithmetic for some $\lambda > 0$ and $M := \mathbb{R}$ if otherwise. Denote T'_k the moments of jumps over $\{\inf B, \sup B\}$; $H'_k := S'_{T_k}$ the overshoots; and put $\kappa' := \min\{k \geq 1 : T'_k \geq \tau_B\}$.

5. General sets

Let M be the state space of the random walk, that is $M := \lambda\mathbb{Z}$ if the walk is λ-arithmetic for some $\lambda > 0$ and $M := \mathbb{R}$ if otherwise. Denote T'_k the moments of jumps over $\{\inf B, \sup B\}$; $H'_k := S'_{T_k}$ the overshoots; and put $\kappa' := \min\{k \geq 1 : T'_k \geq \tau_B\}$.

Theorem 2

Assume that $\mathbb{E}X_1 = 0$, $\mathbb{E}X_1^2 := \sigma^2 \in (0, \infty)$, and B is a bounded Borel set with the non-empty $Int_M(B)$. Then for any $x \in M$,

$$p'_n(x) \sim \frac{\sqrt{2} V_B(x)}{\sigma\sqrt{\pi n}}, \quad V_B(x) := \mathbb{E}_x\left[\sum_{i=1}^{\kappa'} \left|H'_i - H'_{i-1}\right| \mathbb{1}_{\{H'_{i-1} \notin Conv(B)\}}\right].$$

Moreover, this holds uniformly for $x = o(\sqrt{n})$. It is true that $0 < V_B(x) < \infty$ for $x \notin Conv(B)$.

Clearly, $V_{(-d,d)}(x) = V_d(x)$.

6. Heuristics

1. It costs to start at $Conv(B) \setminus B$ and exit from it avoiding B.
2. It costs exponentially in time to stay within $Conv(B)$ so the time spent there is negligible.
3. The rest is as in the basic case.

7. Conditional functional limit theorem

Define $\hat{S}_n(t)$: for $t = k/n$ with a $k \in \mathbb{N}$ put $\hat{S}_n(k/n) := S_k/(\sigma\sqrt{n})$, and define the other values by linear interpolation.

Theorem 3

Under assumptions of Thm 2, for any $x \in M$ such that $V_B(x) > 0$,

$$Law_x(\hat{S}_n(\cdot)|\tau_B > n) \stackrel{\mathcal{D}}{\to} Law(\rho W_+) \qquad in\ C[0,1],$$

where W_+ is a Brownian meander, ρ is a r.v. independent of W_+ with the distribution given by $\mathbb{P}(\rho = \pm 1) = \frac{1}{2} \pm \frac{x - \mathbb{E}_x S_{\tau_B}}{2V_B(x)}$.

7. Conditional functional limit theorem

Define $\hat{S}_n(t)$: for $t = k/n$ with a $k \in \mathbb{N}$ put $\hat{S}_n(k/n) := S_k/(\sigma\sqrt{n})$, and define the other values by linear interpolation.

Theorem 3

Under assumptions of Thm 2, for any $x \in M$ such that $V_B(x) > 0$,

$$Law_x(\hat{S}_n(\cdot)|\tau_B > n) \stackrel{\mathcal{D}}{\to} Law(\rho W_+) \qquad in\ C[0,1],$$

where W_+ is a Brownian meander, ρ is a r.v. independent of W_+ with the distribution given by $\mathbb{P}(\rho = \pm 1) = \frac{1}{2} \pm \frac{x - \mathbb{E}_x S_{\tau_B}}{2V_B(x)}$.
Moreover, $\mathbb{P}_x(T'_{\nu_n} \le b_n|\tau_B > n) \to 1$ for any $b_n \to \infty$, where $\nu_n := \max\{k \ge 0 : T'_k \le n\}$.

The later means that the conditional distributions of the moment of the last jump T'_{ν_n} over the edges of B are tight.
For integer-valued asymptotically α-stable walks ($1 \le \alpha \le 2$) the weak convergence was proved by Belkin '72.

II. The largest gap problem

Define the largest gap (maximal spacing) within the range of S_n:

$$Gap(\{S_k\}_{k\geq 1}^n) := G_n := \max_{1\leq k\leq n-1}\big(S_{(k+1,n)} - S_{(k,n)}\big),$$

where $m_n := S_{(1,n)} \leq S_{(2,n)} \leq \cdots \leq S_{(n,n)} =: M_n$ denote the elements of $S_1, \dots, S_n$ arranged in the weakly ascending order.

Motivation: persistence of iterated random walks $Z_n = Y(|S_n|)$, where $Y(t)$ is a centred Lévy process independent with S_n. Considered by Baumgarten '11, V. '12. Is it true that

$$\mathbb{P}(Z_k > 0, k = 1, \dots, n) \asymp \mathbb{P}(Y(t) > 0, m_n \leq t \leq M_n)?$$

II. The largest gap problem

Define the largest gap (maximal spacing) within the range of S_n:

$$Gap(\{S_k\}_{k\geq 1}^n) := G_n := \max_{1\leq k\leq n-1}\big(S_{(k+1,n)} - S_{(k,n)}\big),$$

where $m_n := S_{(1,n)} \leq S_{(2,n)} \leq \cdots \leq S_{(n,n)} =: M_n$ denote the elements of $S_1, \dots, S_n$ arranged in the weakly ascending order.

Motivation: persistence of iterated random walks $Z_n = Y(|S_n|)$, where $Y(t)$ is a centred Lévy process independent with S_n. Considered by Baumgarten '11, V. '12. Is it true that

$$\mathbb{P}(Z_k > 0, k = 1, \dots, n) \asymp \mathbb{P}(Y(t) > 0, m_n \leq t \leq M_n)?$$

The closest result by Borodin '81: for any aperiodic random walk in $\mathbb{Z}$, the number $E_n := (M_n - m_n) - \#(\{S_k\}_{k=1}^n) + 1$ of non-visited sites within the range satisfies $\frac{E_n}{\sqrt{n}} \xrightarrow{\mathbb{P}} 0$ (under $\mathbb{E}X_1 = 0,\ Var(X_1) < \infty$).

2. The order of G_n.

Proposition (Ding, Peres, V.)

If $\mathbb{E}X_1 = 0$, $Var(X_1) < \infty$, then the family $Law(G_n)_{n\geq 1}$ is tight.

Proof: Notice that for any $h > 0$,

$$\{G_n \geq 2h\} = \bigcup_{k=1}^{n} \Big\{ S_i \notin (S_k, S_k + 2h), i = 1, \dots, n; S_k < M_n \Big\}.$$

By splitting the trajectory at S_k and reversing time for the part $S_1, \dots, S_{k-1}$, we obtain

$$\mathbb{P}(G_n \geq 2h) \leq \dots$$

$$\leq \frac{2}{\sigma^2 \pi} \sum_{k=1}^{n} \frac{\big(V_h(h) - U_h(h)\big) V_h(-h) + \big(V_h(-h) - U_h(-h)\big) V_h(h) + o(1)}{\sqrt{k(n-k+1)}}.$$

3. The limit theorems for G_n and E_n.

Theorem 4

Let S_n be a random walk with $\mathbb{E}X_1 = 0$, $\mathbb{E}X_1^2 < \infty$. For any sequence $b_n \to \infty$ such that $b_n = o(n)$ it holds that

$$G_n^{Int} := \max_{b_n \le k \le n-b_n} (S_{(k+1,n)} - S_{(k,n)}) \xrightarrow{\mathbb{P}} u,$$

where $u = \lambda$ if the walk is λ-arithmetic and $u = 0$ if o/w, and

$$G_n^{Ext} := \max_{k \in [1,b_n] \cup [n-b_n, n-1]} (S_{(k+1,n)} - S_{(k,n)}) \xrightarrow{\mathcal{D}} \max(G^-, G^+),$$

where G^- and G^+ are i.i.d. positive proper random variables. Consequently, $G_n \xrightarrow{\mathcal{D}} \max(G^-, G^+)$.

3. The limit theorems for G_n and E_n.

Theorem 4

Let S_n be a random walk with $\mathbb{E}X_1 = 0$, $\mathbb{E}X_1^2 < \infty$. For any sequence $b_n \to \infty$ such that $b_n = o(n)$ it holds that

$$G_n^{Int} := \max_{b_n \le k \le n-b_n} (S_{(k+1,n)} - S_{(k,n)}) \xrightarrow{\mathbb{P}} u,$$

where $u = \lambda$ if the walk is λ-arithmetic and $u = 0$ if o/w, and

$$G_n^{Ext} := \max_{k \in [1,b_n] \cup [n-b_n, n-1]} (S_{(k+1,n)} - S_{(k,n)}) \xrightarrow{\mathcal{D}} \max(G^-, G^+),$$

where G^- and G^+ are i.i.d. positive proper random variables.
Consequently, $G_n \xrightarrow{\mathcal{D}} \max(G^-, G^+)$.
Moreover, if the walk is λ-arithmetic, then

$$E_n \xrightarrow{\mathcal{D}} E^- + E^+,$$

where E^- and E^+ are i.i.d. proper random variables.

What are $G_\pm, E_\pm$?

Let $S_n^{\geqslant}$ and $S_n^{<}$ be *independent* Markov chains on $[0,\infty)$ and $(-\infty,0)$, resp., that start at 0 with the transition probabilities

$$\mathbb{P}_x(S_1^{\geqslant} \in dy) = \frac{U_{\geqslant}(y)}{U_{\geqslant}(x)} \mathbb{P}_x(S_1 \in dy), \quad x, y \geq 0,$$

$$\mathbb{P}_x(S_1^{<} \in dy) = \frac{U_{<}(y)}{U_{<}(x)} \mathbb{P}_x(S_1 \in dy), \quad x \geq 0, y < 0.$$

These are the Doob h-transforms of the random walk S_n.

Here $U_{\geqslant}(x) = x - \mathbb{E}_x S_{\tau_{(-\infty,0)}}$ and $U_{<}(x) := \mathbb{E}_x S_{\tau_{[0,\infty)}} - x$.

What are $G_\pm, E_\pm$?

Let $S_n^{\geqslant}$ and $S_n^{<}$ be *independent* Markov chains on $[0,\infty)$ and $(-\infty,0)$, resp., that start at 0 with the transition probabilities

$$\mathbb{P}_x(S_1^{\geqslant} \in dy) = \frac{U_{\geqslant}(y)}{U_{\geqslant}(x)}\mathbb{P}_x(S_1 \in dy), \quad x, y \ge 0,$$

$$\mathbb{P}_x(S_1^{<} \in dy) = \frac{U_{<}(y)}{U_{<}(x)}\mathbb{P}_x(S_1 \in dy), \quad x \ge 0, y < 0.$$

These are the Doob h-transforms of the random walk S_n.

Here $U_{\geqslant}(x) = x - \mathbb{E}_x S_{\tau_{(-\infty,0)}}$ and $U_{<}(x) := \mathbb{E}_x S_{\tau_{[0,\infty)}} - x$.

Let

$$G^+ := Gap(\{S_n^{\geqslant}, -S_n^{<}\}_{n\ge 0}), \quad E^+ := \#(\lambda\mathbb{N} \setminus \{S_n^{\geqslant}, -S_n^{<}\}_{n\ge 0}).$$

What are $G_{\pm}, E_{\pm}$?

Let $S_n^{\geqslant}$ and $S_n^{<}$ be *independent* Markov chains on $[0,\infty)$ and $(-\infty,0)$, resp., that start at 0 with the transition probabilities

$$\mathbb{P}_x(S_1^{\geqslant} \in dy) = \frac{U_{\geqslant}(y)}{U_{\geqslant}(x)}\mathbb{P}_x(S_1 \in dy), \quad x, y \geq 0,$$

$$\mathbb{P}_x(S_1^{<} \in dy) = \frac{U_{<}(y)}{U_{<}(x)}\mathbb{P}_x(S_1 \in dy), \quad x \geq 0, y < 0.$$

These are the Doob h-transforms of the random walk S_n.

Here $U_{\geqslant}(x) = x - \mathbb{E}_x S_{\tau_{(-\infty,0)}}$ and $U_{<}(x) := \mathbb{E}_x S_{\tau_{[0,\infty)}} - x$.

Let

$$G^+ := Gap(\{S_n^{\geqslant}, -S_n^{<}\}_{n\geq 0}), \quad E^+ := \#(\lambda\mathbb{N} \setminus \{S_n^{\geqslant}, -S_n^{<}\}_{n\geq 0}).$$

No gaps in the bulk: the Dvoretzky-Erdös-Kakutani Theorem that a Brownian motion never increases and Theorem 3 (which is needed to approach the edges of the range).

Connection with the local time $L^x(t)$ of a Brownian motion $W(t)$

Denote $M := \max_{0 \le t \le 1} W(t), m := \min_{0 \le t \le 1} W(t)$. Then

$$\mathbb{P}\big(L^x(1) > 0 \text{ for all } m < x < M\big) = 1, \quad \mathbb{P}\big(L^M(1) = L^m(1) = 0\big) = 1.$$

Consequently,

$$\mathbb{P}\Big(\min_{m \le x \le M} L^x(1) > 0\Big) = 1.$$

Connection with the local time $L^x(t)$ of a Brownian motion $W(t)$

Denote $M := \max_{0 \le t \le 1} W(t), m := \min_{0 \le t \le 1} W(t)$. Then

$$\mathbb{P}\big(L^x(1) > 0 \text{ for all } m < x < M\big) = 1, \quad \mathbb{P}\big(L^M(1) = L^m(1) = 0\big) = 1.$$

Consequently,

$$\mathbb{P}\Big(\min_{m \le x \le M} L^x(1) > 0\Big) = 1.$$

However, for any $\varepsilon > 0$,

$$\mathbb{P}\Big(\min_{m+\varepsilon \le x \le M-\varepsilon} L^x(1) > 0\Big) = 1. \tag{1}$$

This does not imply our Theorem 4 since there is no general invariance principle for local times: available only for aperiodic walks in $\mathbb{Z}$ and the walks with $\mathbb{E} \exp(itX_1) \in L^2(\mathbb{R})$ (Borodin '80s). However, for such walks (1) matches our result on G_n^{Int} with $b_n \asymp n$.