J. Eur. Math. Soc. 12, 819–854

DOI 10.4171/JEMS/216

Itai Benjamini · Nathanaël Berestycki

Random paths with bounded local time

Received July 14, 2008 and in revised form March 2, 2009

Abstract. We consider one-dimensional Brownian motion conditioned (in a suitable sense) to have a local time at every point and at every moment bounded by some fixed constant. Our main result shows that a phenomenon of entropic repulsion occurs: that is, this process is ballistic and has an asymptotic velocity approximately 4.58... as high as required by the conditioning (the exact value of this constant involves the first zero of a Bessel function). We also study the random walk case and show that the process is asymptotically ballistic but with an unknown speed.

1. Introduction

The goal of this paper is to describe the macroscopic behaviour of a process which locally behaves like a Brownian motion, but which is constrained to satisfy a global constraint of a self-avoiding nature. Informally speaking, we consider one-dimensional Brownian motion conditioned on the event $\mathcal{E}$ that the local time of the process is bounded by a fixed constant, say 1, at every time and position. The event $\mathcal{E}$ has of course zero probability, so a precise definition is needed—this is deferred to the next section. For the moment, it suffices to say that it is possible to define a probability measure $\mathbb{Q}$ on continuous paths corresponding to this conditioning, which is obtained by a limiting procedure.

From an intuitive point of view, one expects that, conditionally on $\mathcal{E}$, the process will be transient and must in fact escape to infinity with a positive velocity. In fact, one expects the speed to be at least equal to 1, since that is precisely what it means to spend less than one unit of local time per level. This being a very costly behaviour for Brownian motion, it is tempting to believe that the process is not likely to satisfy any constraint that would be even stronger, and hence that the speed of the process will in fact be equal to 1 in the limit.

Our main finding in this paper is that this intuition turns out to be erroneous. To be precise, we obtain the following result.

Theorem 1. $\lim_{t\to\infty} X_t/t = \gamma_0$ *exists in* $\mathbb{Q}$*-probability, and furthermore*

$$\gamma_0 = \frac{3}{1-2j_0^{-2}} = 4.5860\ldots \tag{1}$$

I. Benjamini: Mathematics Department, The Weizmann Institute of Science, POB 26, Rehovot 76100, Israel; e-mail: itai.benjamini@weizmann.ac.il

N. Berestycki: Statistical Laboratory, Cambridge University, Wilberforce Rd. CB3 0WB, Cambridge, United Kingdom; e-mail: N.Berestycki@statslab.cam.ac.uk

where $j_0 = 2.4048\ldots$ *is the first nonnegative zero of the Bessel function* $J_0(x)$ *of the first kind and of order* 0.

A more precise result is stated in Theorem 2 in the next section, after the definitions have been given. In Theorem 5, we obtain a similar result for the corresponding random walk problem: given some $L_0 \geq 2$, a simple symmetric random walk on $\mathbb{Z}$ is conditioned to never visit any site more than L_0 times. Under the limiting measure $\mathbb{Q}$, we show that the particle escapes to infinity with a certain speed $\gamma(L_0)$ and we show that $\gamma(L_0) > 1/L_0$.

We call this phenomenon Brownian entropic repulsion, by analogy with a situation arising in the study of the harmonic crystal which will be described below (in Section 3). Roughly speaking, entropic repulsion describes the fact that the easiest way to achieve a certain global constraint for a random process is to achieve much more than required. Here, this phenomenon arises due to the fact the local time process has wild oscillations, and therefore the process must on average have a small local time if it wants to avoid ever being equal to 1. As discussed in Section 3, the situation in the harmonic crystal is not much different. We also describe other conditionings of Brownian motion where a similar entropic repulsion occurs in the paper [1], and recall some results of that paper later on in Section 3. We expect entropic repulsion to be a general principle in this sort of situations, even though it seems hard to even formalize this idea precisely.

Our techniques are very different in the continuous and the discrete cases. In the continuous case, our main tools are the Ray–Knight theorem and some careful coupling estimates. The existence and uniqueness of the measure $\mathbb{Q}$ is obtained by showing that the approximating sequence forms a Cauchy sequence for a suitable metric (and showing that this implies weak convergence). The value of the speed is obtained through a connection to an eigenvalue problem for the Laplacian in the unit disc of the plane (Sturm–Liouville problem). While it seems possible to adapt these techniques to the discrete case, we have used here a rather different method which we believe sheds additional light on the problem. In particular, the notion of regenerating levels plays a significant role in this proof and we crucially apply the renewal theorem in the spirit of Kesten [10]. This theorem is here viewed as a purely analytic result on sequences of numbers satisfying certain conditions, and is applied to sequences which do not have obvious probabilistic interpretations.

2. Statement of the results

Let Ω be the space of continuous, real-valued functions defined on $[0, \infty)$. We endow Ω with the Skorokhod topology and the Borel σ-field defined by it, and with the Wiener measure $\mathbb{W}$. (We let $\mathbb{W}_x$ be the Wiener measure started at the point $x \in \mathbb{R}$). Let $(X_t, t \geq 0)$ be the canonical (coordinate) process on Ω, and let $L(t, x)$ denote a jointly continuous version of the local times process of X, i.e., $\mathbb{W}$-almost surely for all $x \in \mathbb{R}$ and all $t \geq 0$,

$$L(t, x) = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_0^t \mathbf{1}_{\{|X_s - x| \leq \varepsilon\}}\, ds. \tag{2}$$

(We may, occasionally, write $L(t, x, \omega)$ to make explicit the dependence of $L(t, x)$ upon the path $\omega \in \Omega$.) In particular, X satisfies the *occupation formula*: almost surely, for all

$t \geq 0$ and for all nonnegative Borel functions f,

$$\int_0^t f(X_s)\, ds = \int_{\mathbb{R}} f(x) L(t, x)\, dx. \tag{3}$$

(For this and other basic facts about local times of Brownian motion, we refer the reader to Chapter VI of [15]. The above statement corresponds to Corollary (1.6) in that reference.) For all $a > 0$, let

$$\tau_a = \inf\{t > 0 : X_t \geq a\}.$$

We approximate the event $\mathcal{E}$ described in the introduction by conditioning on what happens up to time τ_a, and let a tend to infinity. Hence, define

$$\mathcal{E}_a := \{\sup_{t \leq \tau_a} \sup_{x \in \mathbb{R}} L(x, t) \leq 1\}. \tag{4}$$

A more precise statement of Theorem 1 follows. Recall that $\gamma_0 = 4.5860\ldots$ is defined by (1).

Theorem 2. *The family of measures* $\{\mathbb{P}_a := \mathbb{W}(\cdot \mid \mathcal{E}_a)\}_{a \geq 0}$ *converges weakly to a measure* $\mathbb{Q}$ *on* Ω *as* $a \to \infty$. *Moreover,* $\lim_{t\to\infty} X_t/t = \gamma_0$ *in* $\mathbb{Q}$*-probability.*

Roughly speaking, the idea for the proof of this theorem is that, by the Ray–Knight theorem, the local times of Brownian motion between the two endpoints 0 and a behave as the square radius of a two-dimensional Brownian motion. Conditioning by the event $\mathcal{E}_a$ amounts to conditioning this Brownian motion to stay within the unit ball. By well-known results due to Pinsky [14] on *metastability*, this has a simple equilibrium distribution, under which the square radius has an average $c < 1$. The asymptotic speed of the process is then given by $\gamma_0 = 1/c$, and thus $\gamma_0 > 1$.

Remark 3. If one requires the local time to be bounded by $C > 0$ rather than 1 in the events $\mathcal{E}$ and $\mathcal{E}_a$, it can be shown that the limiting speed of the process becomes γ_0/C. That is, entropic repulsion makes the particle travel $4.5860\ldots$ as fast as one would expect from the constraint in the conditioning.

Remark 4. Conditioning on the event $\mathcal{E}_a$ drives the process to $+\infty$ because the condition is imposed at time τ_a, the hitting time of $a > 0$. If we replace τ_a in the definition of $\mathcal{E}_a$ by τ_a', where

$$\tau_a' = \inf\{t \geq 0 : |X_t| = a\}$$

then Theorem 2 easily implies that the same statement is true with $\lim_{t\to\infty} X_t/t = \pm\gamma_0$ with probability $1/2$ each.

We now turn to the discrete version of the problem, about which we know both more and less information. As our basic probability space we take $\Omega = \{-1, +1\}^{\mathbb{Z}_+}$. A generic point of Ω is written as $\omega = \{\omega_t\}_{t \geq 0}$. For $\omega \in \Omega$, let

$$S_n(\omega) = \sum_{j=1}^{n} \omega_j, \qquad n = 0, 1, \ldots,$$

be the random walk on $\mathbb{Z}$ associated to ω. For $t \in \mathbb{Z}_+$, $x \in \mathbb{Z}$, define

$$L(t,x) = \sum_{j=1}^{t} \mathbf{1}_{\{S_j = x\}}.$$

Of course, $L(t,x)$ is a function on Ω. Occasionally it will be useful to write $L(t,x,\omega)$ for the value of $L(t,x)$ at the point $\omega \in \Omega$. In fact, $L(t,x,\omega)$ depends only on the first $t+1$ coordinates of ω, so we can also regard it as a function on $\Omega_t := \{-1,+1\}^{t+1}$. If $\omega_t \in \Omega_t$ we shall also use the notation $L(t,x,\omega_t)$ for the value of $L(t,x)$ at this point. Unless otherwise indicated we take $S_0 = 0$. Let us now define the event $\mathcal{B}$ which serves as our constraint: for $r \in \mathbb{Z}_+$, let

$$\mathcal{B}_r = \{L(\tau_r, x) \le L_0 \text{ for all } x\} = \{\omega \in \Omega : L(\tau_r(\omega), x, \omega) \le L_0 \text{ for all } x\}$$

where

$$\tau_r = \inf\{k \ge 1 : S_j = r\}.$$

To formulate our main result for random walks we will need to introduce the notion of "regenerating levels", to borrow from the terminology of random walks in random environments. Define

$$\nu_1 = \inf\{r \ge 0 : S_t > r \text{ for all } t \ge \tau_r + 1\}$$

and recursively, for all $i \ge 2$,

$$\nu_i = \inf\{r > \nu_{i-1} : S_t > r \text{ for all } t \ge \tau_r + 1\}.$$

The levels ν_i are those which are visited only once for a given trajectory ω.

Theorem 5. *The measures $\mathbb{P}(\cdot \mid \mathcal{B}_r)$ converge weakly to a limiting measure $\mathbb{Q}$ as $r \to \infty$. Then for all $j \ge 1$, $\nu_j < \infty$ $\mathbb{Q}$-a.s. Moreover, the random variables $(\nu_{j+1} - \nu_j)_{j \ge 1}$ are i.i.d. and satisfy*

$$E_{\mathbb{Q}}(\nu_{j+1} - \nu_j) < \infty.$$

The portions of the path between two successive renewal levels are also independent. In particular, $\gamma(L_0) = \lim_{k\to\infty} X_k/k$ exists $\mathbb{Q}$-almost surely, and is a nonrandom number satisfying $\gamma(L_0) > 1/L_0$.

3. Related work

3.1. Harmonic crystal with hard wall repulsion

As already mentioned above, the term "entropic repulsion" was introduced to describe a situation arising in the study of the discrete Gaussian free field on a lattice (also known as the harmonic crystal) with hard wall repulsion, which presents some strong analogy to the phenomenon described by Theorems 2 and 5. Indeed, in [3], the following result (among other things) is proved. Let $\Phi_N = (\phi_x)_{x \in V_N}$ be the law of a free field on a box $V_N =$

$\{1, \ldots, N\}^2$ with zero boundary conditions and covariance $\mathrm{cov}(\phi_x, \phi_y) = G_N(x, y)$ (the discrete Green function stopped when the walk reaches the outside of the box). Let D_N be a “nice” domain in the box (essentially, the discrete approximation of a smooth fixed domain in $(0, 1)^2$ away from the boundary, blown up by a factor of N), and let $\Omega^+_{D_N}$ be the event that $\phi_x \geq 0$ for all $x \in D_N$. Then, conditionally on $\Omega^+_{D_N}$, the value of the field ϕ_x is typically of order $\log N$, in the following strong sense: for all $\varepsilon > 0$,

$$\lim_{N\to\infty} \sup_{x\in D_N} \mathbb{P}\Big(\Big|\phi_x - \frac{4}{\pi}\log N\Big| \geq \varepsilon \log N \;\Big|\; \Omega^+_{D_N}\Big) = 0. \tag{5}$$

The intuitive reason for this behaviour is the same as in Theorem 2 above. To simplify, due to the wild oscillations of the free field (or the local time field, in our case), the simplest way to achieve the constraint is a global shift which guarantees that the wild oscillations do not break the constraint.

3.2. *Brownian motion with limited local time*

In [1], we have also studied other conditionings of Brownian motion which favour a self-avoiding behaviour, even though the constraint is much softer than the event $\mathcal{E}$. Namely, we discuss Brownian motion conditioned on the event $\mathcal{K}$ that the growth of the local time at the origin is slower than some function of time $f(t)$ where f is nondecreasing but $f(t)t^{-1/2}$ is nonincreasing. We show that if $\int_1^\infty f(t)t^{-3/2}\,dt < \infty$ then the process is transient. We believe this condition to be sharp. In particular, if $f(t) \sim c\sqrt{t}(\log)^{-\gamma}$ for some $c > 0$ and $\gamma \geq 0$, then the process is transient as soon as $\gamma > 1$. In the regime where $0 \leq \gamma \leq 1$, and where we thus anticipate that the process is recurrent, we nonetheless expect an entropic repulsion phenomenon to occur in the sense that $L_t = o(f(t))$ with high probability for $t \to \infty$.

3.3. *Edwards and Domb–Joyce polymer models*

Finally, the present work is closely related to the field of polymer models. The well-known Domb–Joyce model (and its Brownian analogue—the Edwards model) is a model where simple random walk measure is penalized by a weight exponential in the number of self-intersections. More precisely, given an inverse temperature $\beta > 0$, the Domb–Joyce model is defined by looking at the measure μ_N on nearest-neighbours discrete random paths of length N obtained by setting

$$\mu_N(\omega) = \frac{1}{Z_N}\exp\Big(-\beta \sum_{0\leq i<j\leq N} \mathbf{1}_{\{\omega_i=\omega_j\}}\Big)2^{-N}$$

where Z_N is a normalizing constant. Similarly, the Edwards model (in one dimension) is defined by taking a large $T > 0$ and considering the measure μ_T whose density with respect to the Wiener measure is

$$\frac{d\mu_T}{d\mathbb{W}} = \frac{1}{Z_T}\exp\left(-\beta\int_{\mathbb{R}} L(T,x)^2\,dx\right). \tag{6}$$

where $L(t,x)$ is a jointly continuous version of the local time at time t and position x. It is the limit of the distribution of the position of the endpoint under these measures (and their dependence on β) as N or T tend to infinity which is of interest. The main result on this model, proved in [7], is that X_T is approximately normally distributed with a mean $c(\beta)T$ and variance $\sigma^2(\beta)T$. In the case of the Edwards model, these parameters have simple dependence on β: in fact, the variance parameter $\sigma^2(\beta)$ is independent of the self-repellency strength β, while $c(\beta) = b^*\beta^{1/3}$ for some $0 < b^* < \infty$. However, in the discrete Domb–Joyce model, the dependence on β is largely unknown—it is still an open question to show that $c(\beta)$ is monotone in β. See [8] and the references therein for a very interesting account of the theory. See also the paper [13] for a polymer model related to our work, where explicit calculations on the ballistic behaviour of the process can be done.

We note that both the present work (in the continuous case) and the papers [7, 13] use in a fundamental way the Ray–Knight theorem, as well as (for [7]) a connection to an eigenvalue problem for the Laplacian. However, this is where the analogy stops: while [7] requires many difficult analytical estimates, we only require careful but simple-minded probabilistic coupling estimates. Also, in this paper we discuss the full convergence of the path $(X_s, s \geq 0)$ (in the sense of weak limits of measure on paths) rather than its position at a large time. During the revision of this paper we learnt from the referee that Joseph Najnudel [12] has recently constructed a probability measure on Ω corresponding to the whole process in the setup of the Domb–Joyce polymer model. His techniques are very different from ours. Note also that the discrete case uses entirely different techniques. Finally, we mention that it is very likely that our techniques would yield a central limit theorem for the position of the particle in Theorems 2 and 5. We have not tried to pursue this direction.

A related problem has also been studied by Mörters and Sidorova [11], where they analyse the order of magnitude of the maximal displacement of a random walk conditioned on the p^{th} moment of its local time profile being unusually small, for some $p > 1$. More precisely, let

$$\Lambda_n(p) = \sum_{z\in\mathbb{Z}} L(z,n)^p,$$

where $L(z,n)$ denotes the number of visits by a simple random walk to z by time n. They consider the simple random walk conditioned on the event that $\{\Lambda_n(p) < \varepsilon_n E(\Lambda_n(p))\}$ for some sequence $\varepsilon_n = o(1)$. They are able to show that under this conditioning, there exist constants $c_1, c_2 > 0$ such that

$$c_1 \leq \frac{\max_{1\leq i\leq n}|S_n|}{\sqrt{n}\,\varepsilon_n^{-1/(p-1)}} \leq c_2$$

with high probability as $n \to \infty$. Their result is based on a careful large deviation analysis.

4. Existence and uniqueness of the weak limit

We start the proof of Theorem 2 with the existence and uniqueness of a weak limit for the measures $\mathbb{P}_t := \mathbb{W}(\cdot \mid \mathcal{E}_t)$ as $t \to \infty$, for the Skorokhod topology (we refer the reader to [2] for background on weak convergence). In fact, we are going to prove a stronger statement and show that for the total variation distance on sets measurable with respect to $\mathcal{F}_{\tau_a}$ for a fixed arbitrary $a > 0$, the measures $\mathbb{P}_t$ form a Cauchy sequence. For the convenience of the reader, we first explain precisely what we mean by this and then prove that this implies weak convergence with respect to the Skorokhod topology. The remainder of the section will be devoted to the proof that $\mathbb{P}_t$ is a Cauchy sequence in that sense.

Thus, let $a > 0$, and recall that $\tau_a = \inf\{s > 0 : X_s \geq a\}$. For probability measures μ, ν on $(\Omega, \mathcal{F}_{\tau_a})$ define

$$d_a(\mu, \nu) := \sup_{A \in \mathcal{F}_{\tau_a}} |\mu(A) - \nu(A)|. \tag{7}$$

Lemma 6. *Let $\{\mu_t\}_{t \geq 0}$ be a sequence of probability measures on $\mathcal{F}$ such that for every a, the restrictions of μ_t to $\mathcal{F}_{\tau_a}$ form a Cauchy sequence for the distance d_a, i.e., for every $\varepsilon > 0$, there exists t_0 such that for all $s, t \geq t_0$,*

$$d_a(\mu_t, \mu_s) \leq \varepsilon. \tag{8}$$

Assume also that for all $A > 0$ fixed,

$$\lim_{b \to \infty} \limsup_{t \to \infty} \mu_t(\sup_{s < A} |X(s)| > b) = 0. \tag{9}$$

Then there exists a probability measure μ on $(\Omega, \mathcal{F})$ such that $\mu_t \to \mu$ weakly as $t \to \infty$ for the Skorokhod topology on Ω.

Proof. The proof is mostly routine manipulations, so we content ourselves with outlining it. The bottom line is that convergence in total variation distance is typically much stronger than weak convergence. Fix $A \in \mathcal{F}_{\tau_a}$. Then by (8) and (7), we find that $\mu_t(A)$ is a Cauchy sequence, so has a limit $\mu^a(A)$ as $t \to \infty$. It is easy to check that $\mu^a(A)$ is a probability measure on $(\Omega, \mathcal{F})$ (σ-additivity follows from the uniformity over all sets in (7)). Moreover, for every $A \in \mathcal{F}_{\tau_a}$, $\mu_t(A) \to \mu^a(A)$ as $t \to \infty$. From this it is trivial to check that μ^a satisfies the conditions of Kolmogorov's extension theorem, and thus we may define a unique measure μ such that for all $a > 0$,

$$\mu_t(A) \xrightarrow{t \to \infty} \mu(A) \quad \text{for all } A \in \mathcal{F}_{\tau_a}. \tag{10}$$

While it does not seem *a priori* easy to extend (10) to all sets $A \in \mathcal{F}$ such that $\mu(\partial A) = 0$, where ∂A is the boundary of the set A with respect to the Skorokhod topology, we claim that it follows easily from (8) that $\{\mu_t\}_{t \geq 0}$ is a tight family. There are two conditions to verify, of which the first one (nonexplosion in finite time) is part of the assumption on μ_t (see (9)). The second condition to verify is: for all $A > 0$, and for each $\eta > 0$,

$$\lim_{\varepsilon \downarrow 0} \limsup_{t \to \infty} \mu_t\big(\sup\{|X(s') - X(s'')| : 0 \leq s', s'' < A,\ |s' - s''| < \varepsilon\} > \eta\big) = 0. \tag{11}$$

First observe that by (9) and (10), for all $s > 0$ and $E \in \mathcal{F}_s$, we also have

$$\mu_t(E) \to \mu(E) \tag{12}$$

uniformly in $E \in \mathcal{F}_t$ as $t \to \infty$. Therefore, fix $\delta > 0$ and let t_0 be such that for all $E \in \mathcal{F}_A$, $|\mu_t(E) - \mu_s(E)| \le \delta$ for all $t, s \ge t_0$. Thus for $t \ge t_0$, and $E_{\eta,\varepsilon}$ the event in (11), we have

$$|\mu_t(E_{\eta,\varepsilon}) - \mu_{t_0}(E_{\eta,\varepsilon})| \le \delta,$$

from which it follows that

$$\limsup_{t\to\infty} \mu_t(E_{\eta,\varepsilon}) \le \delta + \mu_{t_0}(E_{\eta,\varepsilon}).$$

Now, since μ_{t_0} is the law of a continuous process, $\lim_{\varepsilon\to 0} \mu_{t_0}(E_{\eta,\varepsilon}) = 0$. Therefore,

$$\lim_{\varepsilon\downarrow 0} \limsup_{t\to\infty} \mu_t(E_{\eta,\varepsilon}) \le \delta$$

where $\delta > 0$ is arbitrary. Then (11) follows by letting $\delta \to 0$. Therefore, $\{\mu_t\}_{t\ge 0}$ forms a tight family, and so there exists some weak subsequential limit. On the other hand, by (12) this limit must be μ since the finite-dimensional marginal distributions are specified by events of the form $E \in \mathcal{F}_t$ for some finite $t > 0$. Since the weak subsequential limit is unique and we have proved tightness, we conclude that $\mu_t \to \mu$ weakly as $t \to \infty$, for the Skorokhod topology. $\square$

For the proof that there exists a weak limit to the sequence $\mathbb{W}(\cdot \mid \mathcal{E}_a)$ as $a \to \infty$, we will use Lemma 6. It turns out that (9) is very easy to verify, and the core of the proof is to check (8). Crucial to this proof is the Ray–Knight theorem; we start by reminding the reader the statement of this result, as can be found in [15, Chapter XI.2], or [16, VI. (52.1)] for the formulation we use here.

A *square Bessel process* of dimension $\delta \ge 0$ is the unique strong solution to the stochastic differential equation

$$Z_t = z_0 + 2\int_0^t \sqrt{|Z_s|}\, dB_s + \delta t, \qquad z_0 \ge 0. \tag{13}$$

In the special case where $\delta = 0$ this process is known as the *Feller diffusion*. When δ is an integer ≥ 1, Z can be interpreted as the square Euclidean norm of a δ-dimensional Brownian motion.

Let $(B_t, t \ge 0)$ be a one-dimensional Brownian motion with joint local time process $\{L(t,x)\}_{t\ge 0,\, x\in\mathbb{R}}$, and let $\tau_a = \inf\{t \ge 0 : B_t = a\}$ be the hitting time of a fixed level $a > 0$.

Theorem 7 (Ray, Knight). *For all $a > 0$, the law of $L(\tau_a, a - x)$ is specified by:*

1. $\{L(\tau_a, a-x)\}_{0\le x\le a}$ *is a square Bessel process of dimension* 2*, started at* 0.
2. *Conditionally given $L(\tau_a, 0) = z_0 \ge 0$, $\{L(\tau_a, -x)\}_{x\ge 0}$ is a Feller diffusion started at z_0 and is independent of $\{L(\tau_a, a-x)\}_{0\le x\le a}$.*

We now state a lemma which allows us to compare different constraints on a 2-dimensional square Bessel process, which will be used repeatedly throughout the proof. It should be noted that in general there is no known way to compare the effect of two different constraints, even when one is intuitively stronger than the other. Lemma 8 shows however that making comparisons is possible when, in some sense, the constraint only deals with the position of the process.

Let $(Y_t, t \geq 0)$ denote a square Bessel process of dimension 2. We may view Y as a random element of $(\Omega, \mathcal{F})$ under the probability measure $\mathbb{Y}$, which is (as explained above) the law on $(\Omega, \mathcal{F})$ of the squared Euclidean norm of a two-dimensional Brownian motion. As we work with many different processes it will at times be convenient to use a generic symbol P for the underlying probability space of different random processes. The notations $\{X(t)\}_{t\geq 0}$, $\{Y(t)\}_{t\geq 0}$ then serve to differentiate these processes, and from the context it should be clear to which processes they refer.

For a given $T > 0$ and a positive measurable function $f : [0, T] \to [0, \infty)$, let $\{Y^f(t)\}_{t\geq 0}$ denote a version of Y conditionally given $\mathcal{A}(0, x) = \{\omega \in \Omega : \omega_s \leq f(s)$ for all $s \leq T$; and $\omega_0 = x\}$. For $0 \leq u \leq T$, we also define the event

$$\mathcal{A}(u, x) = \{\omega \in \Omega : \omega_s \leq f(u + s) \text{ for all } s \leq T - u; \text{ and } \omega_0 = x\}. \tag{14}$$

Lemma 8. $\{Y^f(t), 0 \leq t \leq T\}$ *is an inhomogeneous diffusion on* $\mathbb{R}_+$ *which satisfies*

$$dY^f(t) = \sqrt{Y^f(t)}dB_t + \{2 - \delta^f(t, Y^f(t))\}dt \tag{15}$$

where $\delta^f(t, y) \geq 0$ *for all* $0 \leq t \leq T$ *and for all* $0 < y < f(t)$. *Moreover, if* g *is another function such that* $g(t) \leq f(t)$ *for all* $0 \leq t \leq T$, *then*

$$\delta^g(t, y) \geq \delta^f(t, y) \tag{16}$$

for all $0 < y < g(t)$. *As a result,* $Y \succeq Y^f \succeq Y^g$, *where* $\succeq$ *stands for stochastic domination.*

Proof. It is a well-known fact the conditioned process Y^f can be realized as an h-transform of the original process Y: more precisely, by Girsanov's theorem, Y^f is an inhomogeneous diffusion having the form (15) where

$$\delta^f(t, y) = -\frac{\partial}{\partial y} \log h(t, y) \tag{17}$$

with

$$h(t, y) = \mathbb{Y}(\mathcal{A}(t, y)). \tag{18}$$

(Details can be found for instance in [16, IV.39] in the case where the process Y is Brownian motion. Generalization to weak solutions of stochastic differential equations presents no difficulty and we do not give the details here.) For the first part of Lemma 8 it thus suffices to prove that

$$\frac{\partial h(t, y)}{\partial y} \leq 0. \tag{19}$$

Let $\varepsilon > 0$, and let $y = x + \varepsilon$. It suffices to prove that for all $\varepsilon > 0$ small enough,

$$\mathbb{Y}(\mathcal{A}(t, x)) \geq \mathbb{Y}(\mathcal{A}(t, y)). \tag{20}$$

We use a coupling technique to prove this. Let Y_1 denote a square Bessel process of dimension 2 started from x and let Y_2 denote an independent square Bessel process of dimension 2, but started from y. Let $t > 0$ and let

$$\hat{f}(s) = f(T - t + s) \quad \text{for all } s \leq \hat{T} := T - t. \tag{21}$$

Let $\tau = \inf\{t > 0 : Y_1(t) = Y_2(t)\}$, and let

$$Y_3(t) = \begin{cases} Y_1(t) & \text{if } t < \tau, \\ Y_2(t) & \text{else.} \end{cases} \tag{22}$$

Then by the strong Markov property, Y_3 has the same distribution as Y_1 and moreover $Y_3(s) \leq Y_2(s)$ for all $s \geq 0$ almost surely. It follows that if

$$Y_2(s) \leq \hat{f}(s) \quad \text{for all } s \leq \hat{T}$$

then automatically

$$Y_3(s) \leq \hat{f}(s) \quad \text{for all } s \leq \hat{T}.$$

The desired (20) follows.

The second part of Lemma 8 is an easy consequence of the first part. Indeed, Y^g can be obtained by conditioning further the process Y^f to stay below the function g. We conclude again by Girsanov's theorem that there exists an additional drift term $\delta^{f,g}(t, y)$ such that

$$dY^g(t) = \sqrt{Y^g(t)}dB_t + \{2 - \delta^f(t, Y^g(t)) - \delta^{f,g}(t, Y^g(t))\}dt \tag{23}$$

and that $\delta^{f,g}(t, y)$ satisfies

$$\delta^{f,g}(t, y) = \frac{\partial}{\partial y} \log h^{f,g}(t, y). \tag{24}$$

This time,

$$h^{f,g}(t, y) = P(Y^{\hat{f}}(s) \in \mathcal{A}'(t, y)) \tag{25}$$

where $\hat{f}$ is defined in (21) and $\mathcal{A}'(t, y)$ has the same definition as $\mathcal{A}(t, y)$ except that f is replaced with g. Since $Y^{\hat{f}}$ is a strong Markov process by the first part, the coupling argument works equally well to show that

$$\frac{\partial}{\partial y} h^{f,g}(t, y) \leq 0. \tag{26}$$

As above, this implies $\delta^{f,g}(t, y) \geq 0$ for all $t \leq T, 0 < y < g(t)$, and thus $\delta^g(t, y) \geq \delta^f(t, y)$. To get the final statement of the lemma, we note that it is easy to show that (15) admits strong and pathwise unique solutions, since the coefficients are locally Lipschitz. From this and Theorem 3.7 in [15], the desired stochastic dominations follow directly. □

We now show that if $\mathbb{P}_t := \mathbb{W}(\cdot \mid \mathcal{E}_t)$ for $t > 0$, then $\mathbb{P}_t$ satisfies the assumptions of Lemma 6. Let $s, t > 0$ with $s < t$ and let $0 < a < s$. First note that if $A \in \mathcal{F}_{\tau_a}$, then by elementary manipulations we have

$$\mathbb{P}_t(A) = \mathbb{P}_s(A)\frac{\mathbb{W}(\mathcal{E}_t \mid A \cap \mathcal{E}_s)}{\mathbb{W}(\mathcal{E}_t \mid \mathcal{E}_s)} \tag{27}$$

so it suffices to prove that the ratio is arbitrarily close to 1, uniformly in s, t large enough, and $A \in \mathcal{F}_{\tau_a}$. We will show the existence of a coupling P between two processes X and Y, having respectively the law of $\mathbb{W}(\cdot \mid \mathcal{E}_s)$ and $\mathbb{W}(\cdot \mid \mathcal{E}_s \cap A)$ such that P-almost surely,

$$L(\tau_s, x, X) = L(\tau_s, x, Y) \quad \text{for all } x \geq a + \Delta \tag{28}$$

where $\Delta < \infty$ P-almost surely, and in fact there exists a random variable Δ^* whose distribution does not depend on any parameter, and such that $\Delta \preceq \Delta^*$, and $\Delta^* < \infty$ almost surely. For the moment, let us admit these facts and see how we proceed with them. Let $(Z_u, u \geq 0)$ be a square Bessel-0 process started at an unspecified point $Z_0 = x \in (0, 1)$. We claim that there exist $C, \alpha > 0$ independent of x such that for all $t > 0$,

$$P(Z_t > 0 \mid \sup_{0 \leq u} Z_u \leq 1;\, Z_0 = x) \leq Ce^{-\alpha t}. \tag{29}$$

This follows easily from the Markov property and the fact in any period of duration 1, Z started from position 1 has a positive probability p_0 to reach zero. If not, then at the next iteration the process is still below 1 and again has a probability bigger than p_0 to die out in the next interval. Using Lemma 8, we conclude that (29) holds with $\alpha = -\log(1 - p_0)$.

To ease notations, let $F_1(u) = 1 - L(\tau_s, t - u, X)$ for all $u \geq 0$, and similarly let $F_2(u) = 1 - L(\tau_s, t - u, Y)$ for all $u \geq 0$. In particular, note that $F_1(u) = F_2(u) = 1$ for all $u \leq t - s$, P-almost surely. Note that our assumption (28) implies that $F_1(u) = F_2(u)$ for all $u \leq t - (a + \Delta)$. If now $(Z_u, u \geq 0)$ is the Ray–Knight diffusion changing dimension at time $u = t - s$, then we have

$$\mathbb{W}(\mathcal{E}_t \mid \mathcal{E}_s) = P(Z_s \leq F_1(s) \text{ for all } s \geq 0)$$

and

$$\mathbb{W}(\mathcal{E}_t \mid A \cap \mathcal{E}_s) = P(Z_s \leq F_2(s) \text{ for all } s \geq 0).$$

Let E_1, E_2 be the two events in the above equations. It follows that if $p := \mathbb{W}(\mathcal{E}_t \mid \mathcal{E}_s) = P(E_1)$ and $q := \mathbb{W}(\mathcal{E}_t \mid A \cap \mathcal{E}_s) = P(E_2)$, we have

$$p = P(E_1; Z_{t-a-\Delta} = 0) + P(E_1; Z_{t-a-\Delta} > 0)$$

while

$$q = P(E_2; Z_{t-a-\Delta} = 0) + P(E_2; Z_{t-a-\Delta} > 0).$$

By definition of Δ, we must have $P(E_1; Z_{t-a-\Delta} = 0) = P(E_2; Z_{t-a-\Delta} = 0)$, so it follows that

$$|p - q| = P(E_1; Z_{t-a-\Delta} > 0) + P(E_2; Z_{t-a-\Delta} > 0).$$

and thus

$$\left|1-\frac{q}{p}\right| = P(Z_{t-a-\Delta} > 0 \,|\, E_1) + \frac{P(E_2;\, Z_{t-a-\Delta} > 0)}{p}. \tag{30}$$

We study the two terms on the right hand side separately. For the first term, we note that by (29) and Lemma 8, we get

$$P(Z_{t-a-\Delta} > 0 \,|\, E_1) \le E(Ce^{-\alpha(s-a-\Delta)_+}) \le C\ E(e^{-\alpha(s-a-\Delta^*)_+})$$

where $x_+ = \sup(x, 0)$ is the positive part of x. Similarly, the second term in (30) satisfies

$$\begin{aligned}\frac{P(E_2;\, Z_{t-a-\Delta} > 0)}{p} &= \frac{P(E_2;\, Z_{t-a-\Delta} > 0)}{q}\,\frac{q}{p} = P(Z_{t-a-\Delta} > 0 \,|\, E_2)\frac{q}{p}\\ &\le CE(e^{-\alpha(s-a-\Delta^*)_+})\frac{q}{p}\end{aligned}$$

by another application of Lemma 8. To put these two things together, define $\varepsilon := CE(e^{-\alpha(s-a-\Delta^*)_+}) < \infty$ and let $x = q/p$. Thus we have proved

$$|1-x| \le \varepsilon + \varepsilon x.$$

Thus $x - 1 \le \varepsilon + \varepsilon x$ and solving this inequality we find $x \le (1+\varepsilon)/(1-\varepsilon) = 1 + 2\varepsilon/(1-\varepsilon)$. Note that by the Lebesgue convergence theorem, $\varepsilon \to 0$ as $s \to \infty$. Thus if s is large enough that $\varepsilon/(1-\varepsilon) \le 2\varepsilon$, we have proved that

$$x \le 1 + 4\varepsilon$$

and a similar lower bound follows without any difficulty. From this and (27), we see that for any $\eta > 0$, there exists $s_0 > a$ large enough that for all $s_0 < s < t$,

$$|\mathbb{P}_t(A) - \mathbb{P}_s(A)| \le \eta$$

for all events $A \in \mathcal{F}_{\tau_a}$. In other words, we have proved that if μ_t is the law of $(X_r, r \le \tau_a)$ under $\mathbb{P}_t$, then

$$d_a(\mu_t, \mu_s) \le \eta$$

for all $s, t \ge s_0$. That is, $\{\mu_t\}_{t\ge 0}$ forms a Cauchy sequence for the total variation distance. Condition (9) is a direct consequence of Lemmas 15 and 12, so the proof is deferred to the next section. Thus, provided (28) holds, μ_t satisfies the assumptions of Lemma 6 and therefore has a (unique) weak limit μ.

We now turn to the proof of (28). This is based on a time-reversal argument and coupling. Note first that it suffices to construct a coupling of $\{L(\tau_s, x, X)\}_{x\in\mathbb{R}}$ and $\{L(\tau_s, x, Y)\}_{x\in\mathbb{R}}$ which achieves (28). Combining the Markov property at time τ_a for X with the Ray–Knight theorem, we find that if $Z_x = L(\tau_s, x, X)$, then conditionally on $\{Z_a = z \in (0,1)\}$, the process $\{Z_{s-x}\}_{0\le x\ge s-a}$ is a square Bessel-2 process conditioned to never exceed 1 and to be at z at time s. In other words, it is a square Bessel bridge of dimension 2 from 0 to z of duration $s-a$, conditioned never to exceed 1 on that interval. There is naturally a similar description for Y: if $Z'_x = \{L(\tau_s, x, Y)\}$, then conditionally on $\{Z'_a = z' \in (0,1)\}$, the process $(Z'_{s-x}, 0 \le x \ge s-a)$ is a square Bessel bridge from 0 to z' in duration $s-a$, conditioned never to exceed 1 during that interval. We can now return time and say that, conditionally on $\{Z_a = z \in (0,1)\}$ (resp. $\{Z'_a = z'\}$), the process $(Z_x, a \le x \ge s)$ (resp. $(Z'_x, a \le x \le s)$) is a square Bessel bridge from z (resp. z') to 0 in duration $s-a$, conditioned never to exceed 1 during that interval. This being an (inhomogeneous) Markov processes, we can couple the processes Z and Z' after the first time (above level a) that they meet. That is, let $z < z' \in (0,1)$ without loss of generality, and let $(Z_x, s \le x \le a)$ and $(Z'_x, s \le x \le a)$ be two independent square Bessel bridges conditioned never to exceed 1, started respectively from z and z'. Consider $\Delta = \inf\{x \ge a : Z_x = Z'_x\}$. Then the process $\hat{Z}$ defined by

$$\hat{Z}_x = \begin{cases} Z'_x & \text{if } x \le \Delta, \\ Z_x & \text{if } x > \Delta, \end{cases}$$

has the same distribution as Z' and satisfies (28). It thus suffices to show that there exists Δ^* independent of z, z', and a, s and t, such that

$$\Delta \preceq \Delta^* \tag{31}$$

and $\Delta^* < \infty$ almost surely. We will show that in any interval of duration 1, the two processes have a positive probability p to meet, independently of anything in their past. This will show the inequality (31) holds with Δ^* a certain geometric random variable. By Lemma 8, Z' is stochastically dominated by an unconditional square Bessel-2 process started from 1, so for any $s \le x \le a$, and any $\eta > 0$,

$$P(\inf_{y\in[x,x+1]} Z'_y \le \eta \mid \sigma(Z'_y, y \le x)) \ge p_1 \quad \text{a.s.}$$

for some $p_1 > 0$ (note that $p_1 = p_1(\eta)$ depends only on η). This provides an upper bound for Z' and it remains to give a similar lower bound for Z. This takes a few more steps: indeed, it is not hard to see that by the second part of Lemma 8, for any $x \in [s,a]$, $(Z_y, x \le y \le x+1)$ dominates stochastically a square Bessel bridge $(b_y, x \le y \le x+1)$ of dimension 2 from 0 to 0 in duration 1, conditioned on the event $E = \{\sup_{x\le y\le x+1} b_y \le 1\}$. This event E has positive probability, p_2 say. It follows that

$$\begin{aligned} P(\sup_{y\in[x,x+1]} Z_y \le \eta \mid \sigma(Z'_y, y \le x)) &\le P(\sup_{x\le y\le x+1} b_y \le \eta \mid E) \quad \text{a.s.} \\ &\le P(b_{x+1/2} \le \eta)/p_2. \end{aligned}$$

Now, as $\eta \to 0$, the right hand side tends to 0, so we can find $\eta > 0$ small enough (and universal) such that the right hand side is smaller than $1/2$ say. Taking the corresponding $p_1(\eta)$, it follows from the above considerations that

$$P(\Delta \le x+1 \mid \Delta \ge x) \ge p_1(\eta)/2,$$

so taking Δ^* a geometric random variable with success probability $p_1(\eta)/2$ gives us what we were looking for. □

5. Ballistic behaviour

We start with the identification of the value of the limiting speed, which is obtained by solving a certain eigenvalue problem for the Laplacian in two dimensions. Here again our main tools are the Ray–Knight theorem and some careful comparisons obtained through coupling arguments.

Let $\mathbb{Y}^T$ be the law of a square Bessel process $(Y_t, t \geq 0)$ of dimension 2, conditioned on $\{\sup_{s\leq T} Y_s \leq 1\}$. The expectation under this probability measure will be denoted by $E_{\mathbb{Y}^T}(X)$ for a random variable $X \geq 0$.

Lemma 9. *We have*

$$\lim_{t\to\infty} \lim_{T\to\infty} E_{\mathbb{Y}^T}(Y_t) = m_0 = \gamma_0^{-1} = \frac{1 - 2j_0^{-2}}{3}. \tag{32}$$

Proof. Step 1. We start by observing that the measure $\mathbb{Y}_0^T$ is the law of $(|Z^T(t)|^2, t \geq 0)$, where Z^T is a 2-dimensional Brownian motion conditioned not to exit the unit disc $\mathbf{D}$ by time T. By a theorem of Pinsky [14], the distribution of $\{Z(t)\}_{t\geq 0}$ converges as $T \to \infty$ to a diffusion $\{Z^\infty(t)\}_{t\geq 0}$, which can be determined explicitly. We will not be interested in the precise form of the generator of Z^∞. However, we will need to focus on the long term behaviour of the process Z^∞. From the same paper, it is known that Z^∞ admits an invariant nontrivial probability measure measure π on $\mathbf{D}$ whose density is equal to

$$\pi(dx) = \frac{1}{C}\varphi(x)^2 dx \tag{33}$$

where φ is the principal eigenfunction associated with the smallest eigenvalue of the operator $L = -\frac{1}{2}\Delta$ with Dirichlet boundary conditions on $\mathbf{D}$, and $C = \int_{\mathbf{D}} \varphi(x)^2\,dx$. (Note that it does not matter how we have normalised φ here.) That is,

$$\begin{cases} \frac{1}{2}\Delta\varphi &= -\lambda\varphi, \\ \varphi|_{\partial\mathbf{D}} &= 0. \end{cases} \tag{34}$$

It is well-known that the problem (34) has solutions only for a discrete set of values $\{\lambda_0 < \lambda_1 < \cdots\}$ where the lowest eigenvalue is simple, i.e., the corresponding eigenspace is one-dimensional, generated by an eigenfunction denoted by φ_0, the principal eigenfunction. Thus $\varphi = \varphi_0$, which is well-known to be rotationally invariant (a good reference at this level of generality is Jost [9, Chapter 9.5]). Hence $\varphi(x)$ takes the same value over the entire circle of radius $0 < r < 1$. We may thus define a function $\phi(r)$ on $(0, 1)$ such that $\phi(r) = \varphi(x)$ for all $x \in \mathbf{D}$ such that $|x| = r$. By the ergodic theorem ([16, V.54]) applied to the diffusion $(Z_t^\infty, t \geq 0)$, it follows that

$$\lim_{t\to\infty} E(|Z_t^\infty|^2) = m_0 := \frac{1}{C}\int_{\mathbf{D}} |x|^2\varphi(x)^2\,dx. \tag{35}$$

Therefore, $\lim_{t\to\infty} \lim_{T\to\infty} E_{\mathbb{Y}_0^T}(Y_t)$ exists and is equal to $m_0 = (1/C)\int_{\mathbf{D}} |x|^2\varphi(x)^2\,dx$.

Step 2. It turns out that this integral can be evaluated explicitly. The principal eigenfunction can be identified explicitly as (see, e.g., Courant and Hilbert [4, (29) in Chapter V])

$$\phi(r) = J_0(j_0 r) \tag{36}$$

where J_0 is the Bessel function of the first kind of order $\nu = 0$, $j_0 = 2.4048\ldots$ is the first nonnegative zero of J_0. Having chosen this normalisation of ϕ, C is given by

$$C = \int_0^1 J_0(j_0 r)^2 2\pi r\, dr. \tag{37}$$

It follows that

$$m_0 = \frac{2\pi \int_0^1 r^3 J_0(j_0 r)^2\, dr}{2\pi \int_0^1 r J_0(j_0 r)^2\, dr} = j_0^{-2} \frac{\int_0^{j_0} x^3 J_0(x)^2\, dx}{\int_0^{j_0} x J_0(x)^2\, dx}. \tag{38}$$

We turn to the following result which can be found in [17, p. 137], known as Schafheitlin's reduction formula: for all $z \geq 0$, and all $\mu \geq 0$,

$$\begin{aligned}(\mu+2)\int_0^z x^{\mu+2} J_0(x)^2\, dx &= -\frac{1}{4}(\mu+1)^3 \int_0^z x^{\mu} J_0(x)^2\, dx \\ &+ \frac{1}{2}\left[x^{\mu+1}\left(x J_0'(x) - \frac{1}{2}(\mu+1) J_0(x)\right)^2 + x^{\mu+1}\left(x^2 + \frac{1}{4}(\mu+1)^2\right) J_0(x)^2\right]_0^z.\end{aligned} \tag{39}$$

Taking $\mu = 1$ and $z = j_0$ and recalling that $J_0(j_0) = 0$, we obtain

$$3\int_0^{j_0} x^3 J_0(x)^2\, dx = -2\int_0^{j_0} x J_0(x)^2\, dx + \frac{1}{2} j_0^4 J_0'(j_0)^2. \tag{40}$$

Thus

$$3\frac{\int_0^{j_0} x^3 J_0(x)^2\, dx}{\int_0^{j_0} x J_0(x)^2\, dx} = -2 + \frac{j_0^4 J_0'(j_0)^2}{2\int_0^{j_0} x J_0(x)^2\, dx}. \tag{41}$$

It also turns out that

$$\int_0^{j_0} x J_0(x)^2\, dx = \frac{1}{2} j_0^2 J_0'(j_0)^2. \tag{42}$$

(This is a consequence of the fact that the Bessel functions are orthonormal for the weight x: this is a classical property which can be found in [17, p. 576] for instance.) Thus, using (38) together with (41) and (42) we obtain

$$m_0 = j_0^{-2}(1/3)[-2 + j_0^2] = (1/3)[1 - 2j_0^{-2}]. \tag{43}$$

This completes the proof of Lemma 9. □

For $0 \leq x < 1$ and $T > 0$, let $\mathbb{Y}_x^T$ denote the law

$$\mathbb{Y}_x^T(\cdot) = \mathbb{Y}(\cdot \mid Y_0 = x; \sup_{0\leq s\leq T} Y_s \leq 1) \tag{44}$$

and let

$$\mathbb{Y}_x^\infty = \lim_{T\to\infty} \mathbb{Y}_x^T \tag{45}$$

be the weak limit of $\mathbb{Y}_x^T$, which may be described with Pinsky's result [14].

Lemma 10. *For any $\varepsilon, \eta > 0$, there exists t_0 such that for all $t \geq t_0$, and for all large enough $T > 0$,*

$$\mathbb{Y}_{1/2}^T\left(\left|\frac{1}{t}\int_0^t Y_s\,ds - m_0\right| > \varepsilon\right) \leq \eta. \tag{46}$$

Proof. Let $t > 0$. As $t \to \infty$, we know by the ergodic theorem for one-dimensional diffusions (Theorem V.53.1 in [16]) and the above calculations that, $\mathbb{Y}_{1/2}^\infty$-almost surely,

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t Y_s\,ds = m_0. \tag{47}$$

Thus this convergence holds in $\mathbb{Y}_{1/2}^\infty$-probability as well, and we may choose t_0 large enough that

$$\mathbb{Y}_{1/2}^\infty\left(\left|\frac{1}{t}\int_0^t Y_s\,ds - m_0\right| > \varepsilon\right) \leq \eta/2 \tag{48}$$

for all $t \geq t_0$. Let us fix any $t \geq t_0$. Since $\mathbb{Y}_{1/2}^T$ converges weakly towards $\mathbb{Y}_{1/2}^\infty$, and since integration over the compact interval $[0, t]$ is a continuous functional, we conclude that

$$\mathbb{Y}_{1/2}^T\left(\int_0^t Y_s\,ds \in B\right) \to \mathbb{Y}_{1/2}^\infty\left(\int_0^t Y_s\,ds \in B\right)$$

for all Borel sets $B \subset \mathbb{R}$, as $T \to \infty$. Taking $B = [(m_0 - \varepsilon)t, (m_0 + \varepsilon)t]$, we may choose T_0 large enough that for all $T \geq T_0$,

$$\left|\mathbb{Y}_{1/2}^\infty\left(\left|\frac{1}{t}\int_0^t Y_s\,ds - m_0\right| > \varepsilon\right) - \mathbb{Y}_{1/2}^T\left(\left|\frac{1}{t}\int_0^t Y_s\,ds - m_0\right| > \varepsilon\right)\right| \leq \eta/2. \tag{49}$$

Combining (48) and (49) gives the result. □

The next step is to extend Lemma 10 to a similar convergence type of result, but where the starting point x is not necessarily equal to 1/2, while keeping the estimates uniform in x.

Lemma 11. *For any $\varepsilon, \eta > 0$, there exists t_0 such that for all $t \geq t_0$, for all $x \in [0, 1)$, and for all T large enough,*

$$\mathbb{Y}_x^T\left(\left|\frac{1}{t}\int_0^t Y_s\,ds - m_0\right| > \varepsilon\right) \leq \eta. \tag{50}$$

Proof. We prove this by coupling. Consider two independent processes Y^1 and Y^2 sampled respectively from $\mathbb{Y}^T_{1/2}$ and $\mathbb{Y}^T_x$. Let $\tau = \tau(x, T) = \inf\{s > 0 : Y^1_s = Y^2_s\}$, and define

$$Y^3_s = Y^2_s \mathbf{1}_{\{s \le \tau\}} + Y^1_s \mathbf{1}_{\{s \ge \tau\}}. \tag{51}$$

It is easy to show that Y^3 has the same distribution as Y^2, i.e., its law is $\mathbb{Y}^T_x$. Moreover, an application of Lemma 8 shows that the random variable τ is bounded above stochastically, uniformly in T and $x \in [0, 1)$. That is, for any η, there exists $t_1 > 0$ such that for all T large enough and for all $x \in [0, 1)$,

$$P(\tau > t_1) \le \eta. \tag{52}$$

Indeed, the coupling time τ is smaller than the meeting time of two independent processes given by an unconditional square Bessel process of dimension 2 started at 1, with the diffusion $\mathbb{Y}^\infty_0$. This meeting time is finite almost surely, which proves (52). Let $\varepsilon, \eta > 0$. If we now choose t large enough that $t_1/t \le \varepsilon$ and $t > t_0$ from Lemma 10, we obtain

$$\begin{aligned}\mathbb{Y}^T_x\left(\left|\frac{1}{t}\int_0^t Y_s\,ds - m_0\right| > 2\varepsilon\right) &\le P(\tau > t_1) + P\left(\left|\frac{1}{t}\int_0^t Y^1_s\,ds - m_0\right| > \varepsilon\right)\\ &\le \eta + \mathbb{Y}^x_{1/2}\left(\left|\frac{1}{t}\int_0^t Y_s\,ds - m_0\right| > 2\varepsilon\right).\end{aligned}$$

Taking the limsup as $T \to \infty$, and using Lemma 10, we obtain

$$\limsup_{T\to\infty} \mathbb{Y}^T_x\left(\left|\frac{1}{t}\int_0^t Y_s\,ds - m_0\right| > 2\varepsilon\right) \le 2\eta \tag{53}$$

for all $t \ge \max(t_0, t_1/\varepsilon)$. Lemma 11 is now easily deduced from (53). □

Our next lemma shows that, given $\mathcal{E}_a$, we are unlikely to spend a large amount of time below 0, and this amount can be controlled uniformly over a. In fact, the lemma states that once we reach a given level we are unlikely to spend more than a certain amount of time z below it.

Lemma 12. *For any $\varepsilon > 0$, there exists $z > 0$ such that for all $a > 0$, and for all $0 \le y < a$,*

$$\mathbb{W}_0\left(\int_{\tau_y}^{\tau_a} \mathbf{1}_{\{X_s \le y\}}\,ds > z \,\middle|\, \mathcal{E}_a\right) \le \varepsilon \tag{54}$$

where $\tau_y = \inf\{s > 0 : X_s = y\}$. Similarly, there is $b > 0$ such that for all $a > b$, and all $y \in (b, a)$,

$$\mathbb{W}_0(\inf_{\tau_y \le s \le \tau_a} X_t < y - b) \le \eta. \tag{55}$$

Proof. For $s \geq 0$, let $\tilde{X}_s = X_{\tau_y+s} - y$ and $\tilde{L}(s,w) = L(\tau_y + s, y + w)$. By the Markov property, it is easy to check that, given $\mathcal{E}_a$ and $\mathcal{F}_{\tau_y}$, the process $\tilde{X}$ has the law $\mathbb{W}_0(\cdot \mid \tilde{\mathcal{E}})$, where

$$\tilde{\mathcal{E}} = \{\tilde{L}(s,w) \leq f(w) \text{ for all } 0 \leq s \leq \tilde{\tau}_{a-y}\} \tag{56}$$

with

$$f(w) := 1 - \tilde{L}(0,w) = 1 - L(\tau_y, y+w). \tag{57}$$

For $s \geq 0$, let

$$Z_s = \tilde{L}(\tilde{\tau}_{a-y}, -s) - \tilde{L}(0,-s) \tag{58}$$

be the local time at level $y - s$ accumulated by $\tilde{X}$ after hitting y. Then note that by the occupation formula,

$$\int_{\tau_y}^{\tau_a} \mathbf{1}_{\{X_s \leq y\}}\, ds = \int_0^{\tilde{\tau}_{y-a}} \mathbf{1}_{\{\tilde{X}_s \leq 0\}}\, ds = \int_0^\infty Z_s\, ds. \tag{59}$$

By the Ray–Knight theorem, given $\tilde{\mathcal{E}}$ and $Z_0 = x \in [0,1)$, $(Z_s, s \geq 0)$ has the law

$$\mathbb{Z}_x^f := \mathbb{Z}_x(\cdot \mid \{Z_w \leq f(w) \text{ for all } w \geq 0\}), \tag{60}$$

where $\mathbb{Z}_x$ denotes the law of a Bessel process of dimension 0 started from $Z_0 = x$, i.e., $\mathbb{Z}_x$ is the Feller diffusion started from x. (Note that the event on the right hand side of (60) is an event of positive probability for any given $x < f(0)$, since Feller diffusions become extinct almost surely.) By Lemma 8 applied to the diffusion Z rather than Y, for any $x < f(0)$, the conditional law $\mathbb{Z}_x^f$ is stochastically dominated by $\mathbb{Z}_x$. Using for instance the branching property of Feller diffusions, this is itself dominated by $\mathbb{Z}_1$, since $x < f(0) \leq 1$. Thus, letting $\mu(dx)$ denote the law on $[0, f(0)]$ of Z_0,

$$\begin{aligned}
\mathbb{W}_0\left(\int_{\tau_y}^{\tau_a} \mathbf{1}_{\{X_s \leq y\}}\, ds > z \,\middle|\, \mathcal{F}_{\tau_y}; \mathcal{E}_a\right) &= \mathbb{W}_0\left(\int_0^{\tilde{\tau}_{y-a}} \mathbf{1}_{\{\tilde{X}_s \leq 0\}}\, ds > z \,\middle|\, \tilde{\mathcal{E}}\right) \\
&= \int_0^1 \mu(dx)\, \mathbb{W}_0\left(\int_0^\infty Z_s\, ds > z \,\middle|\, Z_0 = x; \tilde{\mathcal{E}}\right) \\
&\leq \int_0^1 \mu(dx)\, \mathbb{Z}_x^f\left(\int_0^\infty Z_s\, ds > z\right) \\
&\leq \int_0^1 \mu(dx)\, \mathbb{Z}_1\left(\int_0^\infty Z_s\, ds > z\right) \\
&\leq \mathbb{Z}_1\left(\int_0^\infty Z_s\, ds > z\right).
\end{aligned}$$

Now, under $\mathbb{Z}_1$, $(Z_s, s \geq 0)$ is almost surely continuous and becomes extinct in finite time, thus $\int_0^\infty Z_s\, ds < \infty$ almost surely, and the right hand side in the above inequality can be made arbitrarily small for large enough z. Taking the expectation to average out the conditioning of $\mathcal{F}_{\tau_y}$ finishes the proof of the first part of Lemma 12. The second part (55) also follows from the same method; the details are left to the reader. □

We now show how Lemma 12 can be applied to prove a first piece of the result in Theorem 2: it is shown that if $y < a$ is given (we want to think of y large but fixed, and $a \to \infty$), then given $\mathcal{E}_a$ it has taken no more than approximately $m_0 y$ units of time to reach y.

Lemma 13. *For any $\varepsilon, \eta > 0$ there exists y_0 large enough that if $y \geq y_0$, and for all sufficiently large $a > 0$,*

$$\mathbb{W}_0(\tau_y > m_0 y(1+\varepsilon) \mid \mathcal{E}_a) \leq \eta \tag{61}$$

where $\tau_y = \inf\{s > 0 : X_s = y\}$.

Proof. We start by noticing that for any $z \geq 0$,

$$\begin{aligned}\mathbb{W}_0(\tau_y > m_0 y(1+\varepsilon) \mid \mathcal{E}_a) \leq \mathbb{W}_0\bigg(\int_0^y L(\tau_y, w)\,dw > m_0 y(1+\varepsilon) - z \,\bigg|\, \mathcal{E}_a\bigg)\\ + \mathbb{W}_0\bigg(\int_0^{\tau_a} \mathbf{1}_{\{X_s \leq 0\}}\,ds > z \,\bigg|\, \mathcal{E}_a\bigg).\end{aligned}$$

Thus if we choose z as in Lemma 12 applied to $y = 0$, we have, for any $a > 0$, and for any $y \geq y_1 := 2z/(m_0\varepsilon)$

$$\begin{aligned}\mathbb{W}_0(\tau_y > m_0 y(1+\varepsilon) \mid \mathcal{E}_a) &\leq \eta + \mathbb{W}_0\bigg(\int_0^y L(\tau_y, w)\,dw > m_0 y(1+\varepsilon/2) \,\bigg|\, \mathcal{E}_a\bigg)\\ &\leq \eta + \mathbb{W}_0\bigg(\int_0^y L(\tau_a, w)\,dw > m_0 y(1+\varepsilon/2) \,\bigg|\, \mathcal{E}_a\bigg).\end{aligned} \tag{62}$$

For $w \geq 0$, let $Y_w = L(\tau_a, a-w)$. Under $\mathbb{W}_0$, recall that by the Ray–Knight theorem, $(Y_w, w \geq 0)$ is a strong Markov process which has the law of a square planar Bessel process for $(0 \leq w \leq a)$ and a Feller diffusion for $w \geq a$. Now, conditionally on $\mathcal{E}_a$, and conditionally on $Y_a = x \in (0,1)$, it follows easily from the strong Markov property at time a that $(Y_s, 0 \leq s \leq a)$ has the law of a square planar Bessel bridge conditioned on $\{\sup_{s \leq a} Y_s \leq 1\}$. That is, if we further condition on the position $Y_a = x$, the part of the constraint on Y_w for $w \geq a$ becomes irrelevant.

We now appeal to the following time-reversal argument: let $(Y_s, s \geq 0)$ be a square Bessel bridge of dimension 2 with $Y_0 = 0$ and $Y_a = x$, and let

$$Y_w^{\leftarrow} = Y_{a-w}, \quad 0 \leq w \leq a, \tag{63}$$

be the time-reversed process. Then $(Y_w^{\leftarrow}, 0 \leq w \leq a)$ is itself a square Bessel bridge of dimension 2 with $Y_0^{\leftarrow} = x$ and $Y_a^{\leftarrow} = 0$. (This follows quite easily from the rotational invariance of Brownian motion and from the fact that a Brownian bridge presents the same time-reversibility.) Furthermore, note that by Lemma 8, a square Bessel bridge from x to 0, conditioned on $\{Y_s \leq 1 \text{ for all } s \leq a\}$, can be related to the measure $\mathbb{Y}_x^a$ in the following fashion:

$$\mathbb{Y}_x^a(\cdot \mid Y_a = 0) = \lim_{\delta \to 0} \mathbb{Y}_x^a(\cdot \mid Y_a \leq \delta) \preceq \mathbb{Y}_x^a(\cdot) \tag{64}$$

where $\preceq$ stands for stochastic domination. Therefore, taking $Y_w = L(\tau_a, a - w)$,

$$\begin{aligned}\mathbb{W}_0\biggl(\int_0^y L(\tau_a, w)\,dw > m_0 y(1+\varepsilon/2)\,\bigg|\,\mathcal{E}_a, Y_a = x\biggr)&\\ &= \mathbb{Y}_0^a\biggl(\int_{a-y}^a Y_w\,dw > m_0 y(1+\varepsilon/2)\,\bigg|\,Y_a = x\biggr)\\ &= \mathbb{Y}_x^a\biggl(\int_0^y Y_w\,dw > m_0 y(1+\varepsilon/2)\,\bigg|\,|Y_a = 0\biggr)\\ &\le \mathbb{Y}_x^a\biggl(\frac{1}{y}\int_0^y Y_w\,dw > m_0(1+\varepsilon/2)\biggr). \qquad (65)\end{aligned}$$

By Lemma 11, we may choose y_2 large enough that if $y \ge y_2$ and for all large enough a, the right hand side of (65) is smaller than η. Thus for $y \ge y_1 \vee y_2$, and for all large enough a, we have by (62) and unconditioning on the position Y_a in (65),

$$\mathbb{W}_0(\tau_y > m_0 y(1+\varepsilon)\,|\,\mathcal{E}_a) \le 2\eta \qquad (66)$$

as required. □

We now prove a bound in the other direction for the hitting times of certain levels. To start, we need an a priori bound that says that it is unlikely for $L(\tau_a, 0)$ to be close to 1 when we condition on $\mathcal{E}_a$.

Lemma 14. *For any $\eta > 0$, there is a $\delta > 0$ such that*

$$\mathbb{W}_0(L(\tau_a, 0) \ge 1 - \delta\,|\,\mathcal{E}_a) \le \eta \qquad (67)$$

for all large enough $a > 0$.

Proof. By Lemma 8 and the Ray–Knight theorem, we observe that the random variable $L(\tau_a, 0)$, conditionally given $\mathcal{E}_a$, is stochastically dominated by the squared modulus of a two-dimensional Brownian motion at time a, conditioned to be smaller than 1. However, the modulus at time a is an exponential random variable with mean $\sqrt{a}$, so (67) follows easily. □

Lemma 15. *For any $\varepsilon, \eta > 0$ there exists y_3 large enough that if $y \ge y_3$, and for all sufficiently large $a > 0$,*

$$\mathbb{W}_0(\tau_y < m_0 y(1-\varepsilon)\,|\,\mathcal{E}_a) \le \eta. \qquad (68)$$

Proof. The proof proceeds basically through the same steps as Lemma 13, but there are a few changes. Let z be as in Lemma 12, and let $2z/\varepsilon =: y_4 < y < a$. On the event $E(y, z)$ that X does not spend more than z units of time after τ_y below level y, we get

$$\begin{aligned}E(y,z) \cap \{\tau_y < m_0 y(1-\varepsilon)\} &= E(y,z) \cap \left\{\int_{-\infty}^y L(\tau_y, w)\,dw < m_0 y(1-\varepsilon)\right\}\\ &\subset \left\{\int_0^y L(\tau_a, w)\,dw < m_0 y(1-\varepsilon) + z\right\}\\ &\subset \left\{\int_0^y L(\tau_a, w)\,dw < m_0 y(1-\varepsilon/2)\right\}.\end{aligned}$$

Define $(Y_w, w \geq 0)$ to be, as usual, $Y_w = L(\tau_a, a - w)$, for any $w \geq 0$. Recall that Y is an inhomogeneous diffusion, or more precisely, a square Bessel process of dimension 2 on $[0, a]$, and a Feller diffusion on $[a, \infty)$. Fix $\delta > 0$ as in Lemma 14, and note that by optional stopping, since Z is a $\mathbb{Z}_{1-\delta}$-martingale,

$$\mathbb{Z}_{1-\delta}(\sup_{s>0} Z_s < 1) = \delta.$$

Now, by Lemma 11, we can choose y_0 such that if $y \geq y_0$, for all $x \in (0, 1)$ and all $b > 0$ large enough,

$$\mathbb{Y}_x^b\left(\frac{1}{y}\int_0^y Y_s\,ds < m_0 y(1 - \varepsilon/2)\right) \leq \eta\delta. \tag{69}$$

Therefore,

$$\begin{aligned}
&\mathbb{W}_0\left(\int_0^y L(\tau_a, w)\,dw < m_0 y(1 - \varepsilon/2)\,\middle|\,\mathcal{E}_a\right)\\
&\qquad\leq \eta + \mathbb{W}_0\left(\int_0^y L(\tau_a, w)\,dw < m_0 y(1 - \varepsilon/2);\, Y_a \leq 1 - \delta\,\middle|\,\mathcal{E}_a\right)\\
&\qquad\leq \eta + \frac{1}{\delta}\mathbb{Y}_0^a\left(\int_{a-y}^a Y_s\,ds < m_0 y(1 - \varepsilon/2);\, Y_a \leq 1 - \delta\right)\\
&\qquad\leq \eta + \frac{1}{\delta}\mathbb{Y}_0^a\left(\int_{a-y}^a Y_s\,ds < m_0 y(1 - \varepsilon/2)\right).
\end{aligned} \tag{70}$$

The idea is now to condition upon the position $Y_{a-y} = x$. Conditionally on this event,

$$\mathbb{Y}_0^a\left(\int_{a-y}^a Y_s\,ds < m_0 y(1 - \varepsilon/2)\,\middle|\,Y_{a-y} = x\right) = \mathbb{Y}_x^y\left(\frac{1}{y}\int_0^y Y_s\,ds \leq m_0(1 - \varepsilon/2)\right).$$

However, by Lemma 8, $\mathbb{Y}_0^y \preceq \mathbb{Y}_0^b$ for any $b > y$. Thus

$$\mathbb{Y}_0^a\left(\int_{a-y}^a Y_s\,ds < m_0 y(1-\varepsilon/2)\,\middle|\,Y_{a-y} = x\right) \leq \mathbb{Y}_x^b\left(\frac{1}{y}\int_0^y Y_s\,ds \leq m_0(1-\varepsilon/2)\right) \leq \eta\delta$$

by our choice of $y \geq y_0$ and by taking b sufficiently large that (69) holds. Plugging this into (70), we obtain

$$\mathbb{W}_0\left(\int_0^y L(\tau_a, w)\,dw < m_0 y(1 - \varepsilon/2)\,\middle|\,\mathcal{E}_a\right) \leq \eta + \frac{1}{\delta}\eta\delta = 2\eta.$$

This completes the proof of Lemma 15. □

We are now ready to finish the proof of Theorem 2.

Proof of Theorem 2. The proof is divided into two steps, a lower bound and an upper bound. We start with the lower bound. We want to show that for any $\varepsilon, \eta > 0$, there exists t_3 large enough that for all $t \geq t_3$, and for all $a > 0$ sufficiently large,

$$\mathbb{W}_0(X_t < \gamma_0 t(1 - \varepsilon)\,|\,\mathcal{E}_a) \leq \eta. \tag{71}$$

Indeed, if this holds, then it follows by weak convergence that any subsequential limit $\mathbb{Q}$ of $\mathbb{W}_0(\cdot|\mathcal{E}_a)$ satisfies: for any $\varepsilon, \eta > 0$, there exists t_3 such that for all $t \geq t_3$,

$$\mathbb{Q}(X_t < \gamma_0 t(1-\varepsilon)) \leq \eta \tag{72}$$

because the canonical projection map $X \mapsto X_t$ is a continuous map for the Skorokhod topology. Hence $\mathbb{Q}(X_t/t - \gamma_0 < -\varepsilon) \leq \eta$, and we conclude that

$$\mathbb{Q}(X_t/t - \gamma_0 < -\varepsilon) \xrightarrow[t\to\infty]{} 0,$$

which is, as claimed, the lower bound required for the proof of Theorem 2. Let us thus turn to (72), fix $\varepsilon, \eta > 0$ with $\varepsilon < 1$, and choose y_0 as in Lemma 13. For $t_4 = 4y_0 m_0$ and $t \geq t_4$, let $y = \gamma_0 t(1-\varepsilon/2) \geq y_0$. Thus, for all a sufficiently large,

$$\mathbb{W}_0(\tau_y \leq t(1-\varepsilon/4) \,|\, \mathcal{E}_a) \leq \eta.$$

Having reached level $y = \gamma_0 t(1-\varepsilon/2)$ by time $t(1-\varepsilon/4)$, the only way X_t can be below $\gamma_0(1-\varepsilon)t$ is if X reaches again $\gamma_0 t(1-\varepsilon)$ after time τ_y. By (54) in Lemma 12, if $t \geq t_5 = 4b/\varepsilon$ (where z is as in Lemma 12), then this occurs with probability at most η for all large enough a. Thus we conclude, for $t \geq t_3 := t_4 \vee t_5$, for all large enough a,

$$\mathbb{W}_0(X_t < \gamma_0 t(1-\varepsilon) \,|\, \mathcal{E}_a) \leq 2\eta. \tag{73}$$

This concludes the proof of the lower bound. We now turn to the proof of the upper bound, where we wish to prove that for all $\eta, \varepsilon > 0$, there is t_6 large enough that for all $t \geq t_6$, and all $a > 0$ large enough,

$$\mathbb{W}_0(X_t > \gamma_0 t(1+\varepsilon) \,|\, \mathcal{E}_a) \leq \eta. \tag{74}$$

However, note that the event $\{X_t > \gamma_0 t(1+\varepsilon)\}$ is contained in the event $\{\tau_y \leq t\}$ where $y = \gamma_0 t(1+\varepsilon)$. By Lemma 15, if $y \geq y_3$, in particular if $t \geq t_6 := y_3 m_0$, then it follows

$$\mathbb{W}_0(X_t > \gamma_0 t(1+\varepsilon) \,|\, \mathcal{E}_a) \leq \mathbb{W}_0(\tau_y \leq t \,|\, \mathcal{E}_a) \leq \mathbb{W}\left(\tau_y \leq \frac{y m_0}{1+\varepsilon} \,\middle|\, \mathcal{E}_a\right) \leq \eta,$$

as desired. This completes the proof of Theorem 2. □

6. Random walk with bounded local time

Throughout this section we assume

$$L_0 \geq 2. \tag{75}$$

We need to introduce some notation. Let

$$\tau_k := \inf\{i : S_i = k\} \tag{76}$$

be the first hitting time of $k \geq 0$. We then define $\mathcal{B}_k, \mathcal{B}_k^+$ to be the events

$$\mathcal{B}_k := \mathcal{A}_{\tau_k} = \{L(\tau_k, x, \omega) \leq L_0 \text{ for all } x\}, \tag{77}$$

$$\mathcal{B}_n^+ := \mathcal{B}_n \cap \{S_i > 0 \text{ for } 1 \leq i \leq \tau_n\}. \tag{78}$$

Thus $\mathcal{B}_n^+$ occurs if the sample path minus its endpoints stays strictly between its initial point at 0 and its final point at n. Thus the maximum value of the points is n and this is taken on for the first time at the endpoint and necessarily, the length of the path equals τ_n. Moreover, the sample path through time τ_k visits each value x at most L_0 times. The event $\mathcal{B}_k^+$ will play a major role in our analysis, since it can be interpreted as having a regenerating level immediately at the starting point. We shall make use of the following σ-fields:

$$\mathcal{F}_n = \sigma\{S_i, i \leq n\}, \quad \mathcal{F}_\infty = \bigvee_{n \geq 0} \mathcal{F}_n, \quad \mathcal{G}_k = \mathcal{F}_{\tau_k}.$$

Lemma 16. *There exists some constant $C_3 > 0$ such that*

$$P(\mathcal{B}_k^+) \geq C_3 P(\mathcal{B}_k), \quad k \geq 1. \tag{79}$$

Proof. Fix k and let ρ be the last time before τ_k at which the random walk visits 0, i.e.,

$$\rho = \max\{i < \tau_k : S_i = 0\}.$$

Note that $S_{\tau_k} = k > 0$ for $k \geq 1$. Therefore, $S_i > 0$ for $\rho < i \leq \tau_k$. Consequently, a decomposition with respect to the value of ρ shows that

$$\begin{aligned}
P(\mathcal{B}_k) &= \sum_{j=0}^{\infty} P(\rho = j, \mathcal{B}_k) \\
&\leq \sum_{j=0}^{\infty} P(S_j = 0, L(j, x) \leq L_0 \text{ for all } x, \tau_k > j \\
&\quad \text{and } S_n - S_j > 0 \text{ for } 1 \leq n - j \leq \tau_k - j, L(\tau_k, x) - L(j, x) \leq L_0 \text{ for all } x) \\
&= \sum_{j=0}^{\infty} P(S_j = 0, L(j, x) \leq L_0 \text{ for all } x, j < \tau_k) P(\mathcal{B}_k^+).
\end{aligned} \tag{80}$$

But for any x, on the event $\{S_n = y, L(n, x) \leq L_0 \text{ for all } x\}$ we have

$$\begin{aligned}
&P(L(n + 2L_0 + 2, y) \geq L_0 + 1 \mid S_0, \ldots, S_n) \\
&\quad \geq P(S_{n+2i+1} = y + 1,\ S_{n+2i+2} = S_n = y \text{ for } 0 \leq i \leq L_0) \geq 2^{-L_0 - 1} > 0.
\end{aligned}$$

It follows easily from this that

$$P(L(j, x) \leq L_0 \text{ for all } x) = P(\mathcal{A}_j) \leq C_4 e^{-C_5 j} \tag{81}$$

for some constants $0 < C_i < \infty$. In turn, this implies

$$\sum_{j=0}^{\infty} P(S_j = 0, L(j,x) \le L_0 \text{ for all } x,\ j < \tau_k) \le \sum_{j=0}^{\infty} P(L(j,x) \le L_0 \text{ for all } x) < \infty,$$

so that (79) follows from (80). ◻

We need sharper information about possible weak limits of $P(\cdot \mid \mathcal{B}_r)$. This will be given in the following lemma. We define

$$\mathcal{C}_k := \{S_i > k \text{ for all } i > \tau_k\}. \tag{82}$$

Remark 1. We are going to study weak limit points of the measures $P(\cdot \mid \mathcal{B}_r)$ as $r \to \infty$. Note that each $\tau_n < \infty$ a.s. $[P]$, so conditioning on $\mathcal{B}_r$ is the same as conditioning on $\mathcal{B}_r \cap \{\tau_n < \infty\}$ for any n, including $n = r$, possibly. This does not automatically say that for a limit point Q of $P(\cdot \mid \mathcal{B}_{r_i})$ we have $Q(\tau_n < \infty) = 1$ for all n. In fact this will be false for $n < 0$. But it is correct for $n \ge 0$. Indeed, the case $n = 0$ is trivial, since $\tau_0 = 0$ a.s. $[P]$. For $r > n > 0$,

$$P(\tau_n > t, \mathcal{B}_r) = E(\mathbf{1}_{\{\tau_n > t\}} P(\mathcal{B}_r \mid \mathcal{F}_t)) \le P(\tau_n > t) P(\mathcal{B}_{r-n}). \tag{83}$$

To see this, note that if the walk is at a position $m < n$ at time t, then for $\mathcal{B}_r$ to occur the local time has to be $\le L_0$ as the walk moves from m to r, which is an interval of length at least $r - m \ge r - n$: this implies (83). Therefore, by (86) below,

$$P(\tau_n > t \mid \mathcal{B}_r) \le P(\tau_n > t) \frac{P(\mathcal{B}_{r-n})}{P(\mathcal{B}_r)} \le P(\tau_n > t) 2^n.$$

For fixed $n > 0$ we can make the limsup of the right hand side here as $r \to \infty$ as small as we like by taking t large. Thus $Q(\tau_n = \infty) = 0$ for each $n > 0$.

The following lemma is the first of two crucial steps in the proof of Theorem 5.

Lemma 17. *There exists a constant* $0 \le C_4 < \infty$ *such that*

$$\lim_{t \to \infty} P(\mathcal{B}_t)^{1/t} = e^{-C_4} \tag{84}$$

and for all $t \ge 0$,

$$P(\mathcal{B}_t) \ge e^{-C_4 t}. \tag{85}$$

In addition, for all $s, t \ge 0$,

$$P(\mathcal{B}_t) \le 2^s P(\mathcal{B}_{t+s}). \tag{86}$$

Further,

$$P(\mathcal{B}_n^+) \sim C_6 e^{-C_4 n} \tag{87}$$

for a suitable constant $C_6 > 0$.

Proof. For (84) and (85), we merely have to observe that

$$P(\mathcal{B}_{s+t}) = P(\mathcal{A}_{\tau_{s+t}}) \leq P(\mathcal{A}_{\tau_s})P(\mathcal{A}_{\tau_t}) = P(\mathcal{B}_s)P(\mathcal{B}_t), \tag{88}$$

because if the random walk $\{S_n\}$ reaches level $s+t$ at time τ_{s+t}, with $\sup_x L(\tau_{s+t}, x) \leq L_0$, then the random walk must first reach s at time τ_s with $\sup_x L(\tau_s, x) \leq L_0$ and then the random walk starting at s must reach $s + t$ with $\sup_x[L(\tau_{s+t}, x) - L(\tau_s, x)] \leq L_0$. Thus $P(\mathcal{B}_t)$ forms a submultiplicative sequence, and it follows that $\lim_{t\to\infty} P(\mathcal{B}_t)^{1/t} = e^{-C_4}$ exists. It is obvious that $C_4 \geq 0$, and from (86), proved below, we get $C_4 \leq \log 2 < \infty$. Moreover it is well-known that by submultiplicativity, $-C_4 = \inf_{t\geq 1} \log P(\mathcal{B}_t)/t$, hence $P(\mathcal{B}_t) \geq e^{-C_4 t}$ for all $t \geq 1$.

As for (86), this follows from the simple fact that (by definition) the random walk arrives at t for the first time at τ_t, so that $S_{\tau_t} = t$. If then the random walk takes one step to the right it arrives for the first time at $t+1$ at time τ_t+1. Moreover, $\sup_x L(\tau_t+1, x) \leq 1 \vee \sup_x L(\tau_t, x)$, because the random walk visits a new point at $\tau_t + 1$. Thus, if $\mathcal{B}_t$ occurred, then also $\mathcal{B}_{t+1}$ occurs in this case. Hence

$$P(\mathcal{B}_{t+1}) \geq P(\mathcal{B}_t)P(S_{\tau_t+1} = S_{\tau_t} + 1) = \frac{1}{2}P(\mathcal{B}_t).$$

Induction on s now yields (86).

The proof of (87) is much more involved. However, it is closely related to Lemma 2 in Kesten [10]. In analogy with the L_n from this reference we introduce the further event $\mathcal{L}_n$ which is roughly speaking the event that $\mathcal{B}_n^+$ occurs (so that 0 is a regeneration level) and there is no other regeneration level between 0 and n. To give the formal definition, we define the shift T_n by

$$(T_n\omega)_j = \omega_{\tau_n+j}.$$

We then take $\mathcal{B}_0^+$ to be the certain event, $\mathcal{L}_0$ the empty event, and $\mathcal{L}_1 = \mathcal{B}_1^+$ the event $\{S_0 = 0, S_1 = 1\}$. Further, for $n \geq 2$,

$$\mathcal{L}_n := \mathcal{B}_n^+ \cap \{\forall k < n,\ T_k\omega \notin \mathcal{B}_{n-k}^+\}. \tag{89}$$

The last property says that a sample path $(\omega_0, \omega_1, \dots, \omega_m)$ in $\mathcal{L}_n$ cannot be decomposed into two pieces $(\omega_0, \dots, \omega_j)$ and $(\omega_j, \dots, \omega_m)$ with the first part minus its endpoint lying strictly to the left of ω_j and the second part lying strictly to the right of ω_j (except for its initial point). The first part in such a decomposition would belong to $\mathcal{B}_j^+$ and the second part would be a translate of a path in $\mathcal{B}_{n-j}^+$.

Of course $\{T_k\omega \in \mathcal{B}_{n-k}^+\}$ is the event that $\mathcal{B}_{n-k}^+$ occurs for the shifted sequence $T_k\omega = (\omega_{\tau_k}, \omega_{\tau_k+1}, \dots)$. Since $\mathcal{B}_{n-k}^+$ depends only on $(\omega_0, \dots, \omega_{\tau_{n-k}})$ we shall occasionally abuse notation and write $(\omega_{\tau_k}, \omega_{\tau_k+1}, \dots, \omega_{\tau_n}) \in \mathcal{B}_{n-k}^+$ instead of $T_k\omega \in \mathcal{B}_{n-k}^+$.

The main step will be to show that

$$P(\mathcal{B}_n^+) = \sum_{j=1}^{n} P(\mathcal{L}_j)P(\mathcal{B}_{n-j}^+), \quad n \geq 1. \tag{90}$$

This relation holds by convention if $n = 1$, so assume $n \geq 2$ and that $\mathcal{B}_n^+$ occurs. Then define k to be minimal so that $\mathcal{B}_k^+ \cap \{T_k\omega \in \mathcal{B}_{n-k}^+\}$ occurs. This minimal index is well defined because the event $\mathcal{B}_n^+ \cap \{T_n\omega \in \mathcal{B}_0^+\} = \mathcal{B}_n^+$ occurs. Of course the minimal index is unique. We claim that for this minimal k the event $\mathcal{L}_k$ occurs. Indeed, note that $\mathcal{B}_k^+$ occurs, so that by the definition (89) with n and k replaced by k and j, if $\mathcal{L}_k$ fails, then it must be that $\{\forall j < k, T_j\omega \notin \mathcal{B}_{k-j}^+\}$ fails, i.e., there is $j < k$ such that $T_j\omega \in \mathcal{B}_{k-j}^+$. Since $\omega \in \mathcal{B}_n^+$, this implies that $T_j\omega \in \mathcal{B}_{n-j}^+$ as well, and it is obvious that $\mathcal{B}_j^+$ must hold as well since $\mathcal{B}_n^+$ holds. This contradicts the minimality of k, and hence $\mathcal{L}_k$ holds. Since $T_k\omega \in \mathcal{B}_{n-k}^+$ by definition, it follows immediately that

$$P(\mathcal{B}_n^+) \leq \sum_{k=1}^{n} P(\mathcal{L}_k \cap \{T_k\omega \in \mathcal{B}_{n-k}^+\}).$$

But $\mathcal{L}_k \in \mathcal{G}_k$, because the occurrence of $\mathcal{L}_k$ depends on $(\omega_0, \dots, \omega_{\tau_k})$ only. (Recall that $\mathcal{G}_k = \mathcal{F}_{\tau_k}$ by definition.) Thus, by the strong Markov property,

$$P(\mathcal{B}_n^+) \leq \sum_{k=1}^{n} P(\mathcal{L}_k) P(T_k\omega \in \mathcal{B}_{n-k}^+) = \sum_{k=1}^{n} P(\mathcal{L}_k) P(\mathcal{B}_{n-k}^+). \tag{91}$$

To prove the opposite inequality fix a $k \in \{1, \dots, n\}$ and assume the following two events occur:

$$\mathcal{L}_k \quad \text{and} \quad \omega' := T_k\omega = (\omega_{\tau_k}, \omega_{\tau_k+1}, \dots) \in \mathcal{B}_{n-k}^+. \tag{92}$$

Then ω is such that

$$1 \leq \omega_\ell \leq k-1 \quad \text{for } 1 \leq \ell < \tau_k,\ \omega_{\tau_k} = k, \tag{93}$$

and

$$k+1 \leq \omega_{\tau_k+\ell} = \omega'_\ell \leq n-1 \quad \text{for } 1 \leq \ell \leq \tau_n - 1. \tag{94}$$

Moreover, if τ'_{n-k} denotes the first hitting time of $n-k$ by the path ω', then

$$\begin{aligned} \sup_x L(\tau_n, x, \omega) &\leq \sup_x L(\tau_k, x, \omega) \vee \sup_x L(\tau'_{n-k}, x, \omega') \vee 1 \\ &= \sup_x L(\tau_k, x, \omega) \vee \sup_x [L(\tau, n, x, \omega) - L(k, x, \omega)] \vee 1 \leq L_0. \end{aligned} \tag{95}$$

Together these properties show that $\omega \in \mathcal{B}_n^+$. Thus the sample sequences for which the events in (92) occur contribute $P(\mathcal{L}_k)P(\mathcal{B}_{n-k}^+)$ to $P(\mathcal{B}_n^+)$. In order to prove

$$P(\mathcal{B}_n^+) \geq \sum_{j=1}^{n} P(\mathcal{L}_j) P(\mathcal{B}_{n-j}^+) \tag{96}$$

we therefore merely have to show that (92) can occur only for one k. To see that this is indeed the case assume that in addition to (92) also

$$\mathcal{L}_j \quad \text{and} \quad \omega'' := T_j\omega = (\omega_{\tau_j}, \omega_{\tau_j+1}, \dots) \in \mathcal{B}_{n-j}^+ \tag{97}$$

occur for some $j \neq k$, $j \in [1, n]$. For the sake of argument let $j < k$. But then, on the one hand, $\mathcal{B}_j^+$ occurs (by definition of $\mathcal{L}_j$ or since $\mathcal{B}_k^+$ occurs) and on the other hand, also

$$T_j \omega \in \mathcal{B}_{k-j}^+ \text{ occurs.} \tag{98}$$

But this contradicts the definition of $\mathcal{L}_k$, so that (92) and (97) cannot hold simultaneously. This, in turn, implies (96) and then finally (90).

We can finally start the proof of (87) proper. Define

$$f_n = e^{C_4 n} P(\mathcal{L}_n) \quad \text{and} \quad u_n = e^{C_4 n} P(\mathcal{B}_n^+).$$

By our conventions just before (89),

$$u_0 = 1, \qquad f_0 = 0, \qquad u_1 = f_1 = (1/2) e^{C_4}.$$

Moreover, by (90) these quantities satisfy the renewal equation

$$u_n = \sum_{j=1}^{n} f_j u_{n-j}, \quad n \geq 1.$$

In addition, by Lemma 16 and (85),

$$u_n = e^{C_4 n} P(\mathcal{B}_n^+) \geq C_3 e^{C_4 n} P(\mathcal{B}_n) \geq C_3 > 0,$$

and $\lim_{n\to\infty} [u_n]^{1/n} = 1$. By the renewal theorem (see, e.g., Feller [6, Theorems 2 and 3 in 12.3]), these facts imply

$$\sum_{j=1}^{\infty} f_j = 1 \quad \text{and} \quad \lim_{n\to\infty} u_n = \frac{1}{\mu},$$

where

$$0 < \mu = \sum_{j=1}^{\infty} n f_n < \infty. \tag{99}$$

Thus,

$$P(\mathcal{B}_n^+) \sim \frac{1}{\mu} e^{-C_4 n}, \tag{100}$$

which proves (87). The finishes the proof of Lemma 17. □

We now move on to the second crucial step in the proof of Theorem 5.

Lemma 18. *The limit*

$$C_5 := \lim_{n\to\infty} \frac{P(\mathcal{B}_n^+)}{P(\mathcal{B}_n)} \text{ exists} \quad \text{and} \quad C_5 \geq C_3 > 0. \tag{101}$$

Also for $\mathcal{E}$ an event in $\mathcal{G}_k$,

$$\lim_{n\to\infty} P(\mathcal{E} \cap \mathcal{C}_k \mid \mathcal{B}_n) = C_5 e^{C_4 k} P(\mathcal{E}, \sup_x L(\tau_k, x) \leq L_0), \tag{102}$$

where $\mathcal{C}_k = \{S_i > k$ for all $i > \tau_k\}$ is the event defined in (82)*. Finally,*

$$\lim_{n\to\infty} e^{C_4 n} P(\mathcal{B}_n) = C_7 \tag{103}$$

exists, where $0 < C_7 < \infty$.

Proof. Let $\mathcal{R} = \{0 < r_1 < r_2 < \cdots\}$ be a subsequence along which the weak limit of $P(\cdot \mid \mathcal{A}_{\tau_k})$ exists, and let $Q(\cdot)$ be the value of this limit. The limit along the subsequence $\mathcal{R}$ will be denoted as $\lim_{r\in\mathcal{R}}$. Without loss of generality we may assume that also $\lim_{r\in\mathcal{R}} P(\mathcal{B}_r^+)/P(\mathcal{B}_r)$ exists (since it is a bounded sequence) and is at least C_3 (by Lemma 16). (Later on we will prove that this limit does not depend on $\mathcal{R}$ and hence $\lim_{r\to\infty} P(\mathcal{B}_r^+)/P(\mathcal{B}_r)$ exists.) Now let $\mathcal{E} \in \mathcal{G}_k$. Then

$$Q(\mathcal{E}\cap\mathcal{C}_k) = Q(\mathcal{E}, S_i > k \text{ for all } i > \tau_k) = \lim_{N\to\infty} Q(\mathcal{E}, S_i > k, \tau_k < i \le \tau_{k+N}). \tag{104}$$

We want to show that this equals

$$e^{C_4 k} P(\mathcal{E}, \sup_x L(\tau_k, x) \le L_0) \lim_{r\in\mathcal{R}} \frac{P(\mathcal{B}_r^+)}{P(\mathcal{B}_r)}. \tag{105}$$

To this end observe first that

$$Q(\mathcal{E}, S_i > k \text{ for } \tau_k < i \le \tau_{k+N}) = \lim_{r\in\mathcal{R}} \frac{P(\mathcal{E}, S_i > k \text{ for } \tau_k < i \le \tau_{k+N}, \mathcal{B}_r)}{P(\mathcal{B}_r)},$$

and secondly that for $r \ge k+N$ (because $S_{\tau_k} = k$)

$$\begin{aligned}
&|P(\mathcal{E}, S_i > k \text{ for } \tau_k < i \le \tau_{k+N}, \mathcal{B}_r) - P(\mathcal{E}, S_i > k \text{ for } \tau_k < i \le \tau_r, \mathcal{B}_r)| \\
&\quad\le P(S_i = k \text{ for some } \tau_{k+N} < i \le \tau_r, \sup_x L(\tau_r, x) \le L_0) \\
&\quad\le P(\text{there exists some } \tau_{k+N} < i \le \tau_r \text{ for which } S_i = k \\
&\qquad\text{and} \sup_x L(i,x) \le L_0 \text{ as well as } \sup_x [L(\tau_r, x) - L(i,x)] \le L_0) \\
&\quad\le P(S_i = k \text{ and} \sup_x L(i,x) \le L_0 \text{ for some } \tau_{k+N} < i \le \tau_r) P(\mathcal{B}_{r-k}) \\
&\quad\le P(S_i = k \text{ and} \sup_x L(i,x) \le L_0 \text{ for some } \tau_{k+N} < i \le \tau_r) 2^k P(\mathcal{B}_r) \\
&\quad\le 2^k P(\sup_x L(\tau_{k+N}, x) \le L_0) P(\mathcal{B}_r) = 2^k P(\mathcal{B}_{k+N}) P(\mathcal{B}_r) \le C_7 2^k e^{-C_4(k+N)} P(\mathcal{B}_r)
\end{aligned}$$

for some constant C_7 independent of k, r (use Lemma 16 and (100) for the last inequality). Consequently, using (104),

$$Q(\mathcal{E}, S_i > k \text{ for all } i > \tau_k) = \lim_{r\in\mathcal{R}} \frac{P(\mathcal{E}, S_i > k, \tau_k < i \le \tau_r, \mathcal{B}_r)}{P(\mathcal{B}_r)}. \tag{106}$$

But if $S_i > k$ for $\tau_k < i \le \tau_r$, then

$$L(\tau_r, x) = \begin{cases} L(\tau_k, x) & \text{if } x \le k, \\ L(\tau_r, x) - L(\tau_k, x) & \text{if } x > k. \end{cases}$$

Therefore (use $\mathcal{E} \in \mathcal{G}_{\tau_k}$ and again $S_{\tau_k} = k$)

$$P(\mathcal{E}, S_i > k, \tau_k < i \le \tau_r, \mathcal{B}_r) = P(\mathcal{E}, \mathcal{B}_k) P(\mathcal{B}^+_{r-k}). \tag{107}$$

Together with (106) and (100) this proves the desired (105).

We next claim that there exist events $\mathcal{E}_k \in \mathcal{G}_k$ such that

$$\begin{aligned}\{\mathcal{E}_k, S_i > k \text{ for all } i > \tau_k\} &= \mathcal{E}_k \cap \mathcal{C}_k \\ &= \{k \text{ is the smallest value of } n \text{ for which } \mathcal{C}_n \text{ occurs}\}.\end{aligned} \tag{108}$$

To see this, set for $j < k$,

$$\mathcal{D}_{j,k} := \{S_i > j \text{ for } \tau_j < i \le \tau_k\}.$$

Then $\mathcal{C}_j \cap \mathcal{C}_k = \mathcal{D}_{j,k} \cap \mathcal{C}_k$ and consequently

$$\bigcup_{0 \le j < k} (\mathcal{C}_j \cap \mathcal{C}_k) = \mathcal{C}_k \cap \bigcup_{0 \le j < k} \mathcal{D}_{j,k}.$$

The right hand side of (108) equals

$$\mathcal{C}_k \setminus \bigcup_{0 \le j < k} (\mathcal{C}_j \cap \mathcal{C}_k) = \mathcal{C}_k \cap \Big[\bigcup_{0 \le j < k} \mathcal{D}_{j,k} \Big]^c.$$

This gives us (108) with $\mathcal{E}_k$ equal to the complement of $\bigcup_{0 \le j < k} \mathcal{D}_{j,k}$.

We can now apply (105) with $\mathcal{E}$ taken equal to $\mathcal{E}_k$, with the result that

$$\begin{aligned} Q(k \text{ is the smallest value of } n \text{ for which } \mathcal{C}_n \text{ occurs}) &= Q(\mathcal{E}_k \cap \mathcal{C}_k) \\ &= e^{C_4 k} P(\mathcal{E}_k, \sup_x L(\tau_k, x) \le L_0) \lim_{r \in \mathcal{R}} \frac{P(\mathcal{B}^+_r)}{P(\mathcal{B}_r)}. \end{aligned} \tag{109}$$

Finally we shall show that

$$\begin{aligned} \sum_{k=0}^{\infty} Q(k &\text{ is the smallest value of } n \text{ for which } \mathcal{C}_n \text{ occurs}) \\ &= Q(\mathcal{C}_k \text{ occurs for some } k \ge 0) = 1. \end{aligned} \tag{110}$$

From this and (109) we can conclude that

$$C_5 := \lim_{r \in \mathcal{R}} \frac{P(\mathcal{B}^+_r)}{P(\mathcal{B}_r)} \tag{111}$$

exists, is independent of $\mathcal{R}$, and $\ge C_3$ by virtue of Lemma 16. In view of (106) this will also show that for all $\mathcal{E} \in \mathcal{G}_k$ the full limit

$$\lim_{n \to \infty} P(\mathcal{E} \cap \mathcal{C}_k \mid \mathcal{B}_n) = C_5 e^{C_4 k} P(\mathcal{E}_k, \sup_x L(\tau_k, x) \le L_0)$$

exists, and has the value given in (102). Also (103) follows from (87) and (111).

It remains to prove (110). To this end we want to show that $Q(\mathcal{C}_k \mid \mathcal{G}_k)$ is bounded from below. To prove this we note that for each fixed k, any element of $\mathcal{G}_k$ is up to Q-null sets a finite or countable disjoint union of sets of the form

$$\mathcal{H}(\boldsymbol{\eta}) = \{S_i = \omega_i = \eta_i,\ 0 \le i \le m\},$$

where $m < \infty$ and $\boldsymbol{\eta} = (\eta_0, \dots, \eta_m)$ runs over the sequences which satisfy

$$\eta_0 = 0, \quad \eta_i - \eta_{i-1} = \pm 1,\ 1 \le i \le m, \quad \eta_j < \eta_m = k, 0 < j < m. \tag{112}$$

(Note that the requirements on η in (112) are such that $\tau_k = m$ for any sample point with $(S_0, \dots, S_m) = \boldsymbol{\eta}$. We can restrict ourselves to finite m, because $Q(\tau_k = \infty) = 0$ by Remark 1.) Now, as before, for any such $\boldsymbol{\eta}$,

$$\frac{Q(\mathcal{H}(\boldsymbol{\eta}), \mathcal{C}_k)}{Q(\mathcal{H}(\boldsymbol{\eta}))} = \lim_{N\to\infty} \lim_{r \in \mathcal{R}} \frac{P(\mathcal{H}(\boldsymbol{\eta}), S_i > k, \tau_k = m < i \le \tau_{k+N}, \sup_x L(\tau_r, x) \le L_0)}{P(\mathcal{H}(\boldsymbol{\eta}), \sup_x L(\tau_r, x) \le L_0)}. \tag{113}$$

This time we use that the denominator on the right hand side here is at most

$$P(\mathcal{H}(\boldsymbol{\eta}), \sup_x L(\tau_k, x) \le L_0) P(\mathcal{B}_{r-k})$$

(compare (107)). As in the lines following (106) the numerator on the right hand side of (113) is bounded below by

$$\begin{aligned} &P(\mathcal{H}(\boldsymbol{\eta}), \sup_x L(\tau_k, x) \le L_0) P(\mathcal{D}_{k,k+N}, \sup_x [L(\tau_r, x) - L(\tau_k, x)] \le L_0) \\ &\quad = P(\mathcal{H}(\boldsymbol{\eta}), \sup_x L(\tau_k, x) \le L_0) P(\mathcal{B}^+_{r-k}) \ge C_3 P(\mathcal{H}(\boldsymbol{\eta}), \sup_x L(\tau_k, x) \le L_0) P(\mathcal{B}_{r-k}). \end{aligned}$$

It follows from these estimates that

$$\frac{Q(\mathcal{H}(\boldsymbol{\eta}), \mathcal{C}_k)}{Q(\mathcal{H}(\boldsymbol{\eta}))} \ge C_3.$$

Since this holds for all atoms η of $\mathcal{G}_k$ we conclude that

$$Q(\mathcal{C}_k \mid \mathcal{G}_k) \ge C_3. \tag{114}$$

The relation (110) is a simple consequence of (114) and the martingale convergence theorem. Indeed, set

$$Y_N = \mathbf{1}_{\{\mathcal{C}_k \text{ occurs for some } k \ge N\}}.$$

Then, if we write E^Q for expectation with respect to Q, we have, for each fixed N,

$$\lim_{k\to\infty} E^Q(Y_N \mid \mathcal{G}_k) = Y_N \quad \text{a.s. } [Q].$$

On the other hand, for $k \ge N$, $E^Q(Y_N \mid \mathcal{G}_k) \ge Q(\mathcal{C}_k \mid \mathcal{G}_k) \ge C_3$, from which we deduce that

$$Y_N \ge C_3 \quad \text{a.s. } [Q].$$

Thus $Q(Y_N = 1) = 1$ and (110) holds. This finishes the proof of Lemma 18. □

Lemma 19. *The sequence of measures $P(\cdot \mid \mathcal{B}_n)$ converges weakly to a limit measure Q.*

Proof. Because the walk is nearest-neighbour, it is always the case that $P(\cdot \mid \mathcal{B}_n)$ is tight: it thus suffices to prove uniqueness of the weak subsequential limits. Thus, let Q be a weak limit along the subsequence $\mathcal{R}$. Let $\nu_0 = 0$ and let $0 \le \nu_1 < \nu_2 < \cdots$ be the successive values of ν for which $\mathcal{C}_\nu$ occurs. (110) shows that $\nu_1 < \infty$ a.s. $[Q]$, but the proof of (110) shows that all ν_i are a.s. $[Q]$ finite. From this we will see that $Q(\mathcal{E}) = \lim_{n\to\infty} P(\mathcal{E} \mid \mathcal{B}_n)$, with the limit taken along the sequence of all integers, for any cylinder set $\mathcal{E}$. Indeed, let $\mathcal{E} \in \mathcal{F}_t$. Since $\tau_t \ge t$ (because $|S_{i+1} - S_i| \le 1$), we have $\mathcal{F}_t \subset \mathcal{G}_t$, and so $\mathcal{E} \in \mathcal{G}_t$. Now let ρ be the first $\nu_i \ge t$. Then $\mathcal{E} = \bigcup_{s\ge t}[\mathcal{E} \cap \{\rho = s\}]$ and $\mathcal{E} \cap \{\rho = s\} = \mathcal{E}_s \cap \mathcal{C}_s$ for some $\mathcal{E}_s \in \mathcal{G}_s$ (as in (108)). Consequently,

$$\lim_{n\to\infty} P(\mathcal{E} \cap \{\rho = s\} \mid \mathcal{B}_n) \text{ exists (by (102)).}$$

Also,

$$\left| P(\mathcal{E} \mid \mathcal{B}_n) - \sum_{s=t}^{s=t+N} P(\mathcal{E} \cap \{\rho = s\} \mid \mathcal{B}_n) \right| \le P(\rho > t + N \mid \mathcal{B}_n).$$

Finally,

$$\lim_{N\to\infty} \limsup_{n\to\infty} P(\rho > t + N \mid \mathcal{B}_n) = 0, \tag{115}$$

because if this fails, then (by the monotonicity in N) there exists a sequence $\mathcal{R} = \{r_1 < r_2 < \cdots\}$ and an $\varepsilon > 0$ such that

$$Q(\rho \le t + N) \le 1 - \varepsilon \quad \text{for all } N,$$

where Q is the weak limit of $P(\cdot \mid \mathcal{B}_{r_i})$. But we have just seen that $\rho < \infty$ a.s. $[Q]$, so that (115) must hold. But then

$$\lim_{n\to\infty} P(\mathcal{E} \mid \mathcal{B}_n) = \sum_{s=t}^{\infty} \lim_{n\to\infty} P(\mathcal{E} \cap \{\rho = s\} \mid \mathcal{B}_n).$$

This proves Lemma 19. □

From now on Q will be the (weak) limit of the probability measures $P(\cdot \mid \mathcal{B}_n)$ on Ω. Since $S_0 = 0$ and $S_{n+1} - S_n = \pm 1$ with P-probability 1, it is also the case that

$$Q(S_0 = 0) = 1 \quad \text{and} \quad Q(S_{n+1} - S_n = \pm 1) = 1. \tag{116}$$

Also

$$Q(S_n = y \text{ for more than } L_0 \text{ values of } n) = 0, \tag{117}$$

because for each fixed n and all $r > n$,

$$P((L_0 + 1)^{\text{th}} \text{ visit of } S \text{ to } y \text{ is at time } n \mid \mathcal{B}_r) = 0.$$

We remind the reader that $\mathcal{C}_k$ is defined in (82). We now come to our main result, which describes the structure of Q and is a more precise statement than Theorem 5. Define $\sigma_0 = 0$,

$$\sigma_1 := \inf\{\tau_\ell : \mathcal{C}_\ell \text{ occurs}\}, \quad \sigma_{j+1} := \inf\{\tau_\ell > \sigma_j : \mathcal{C}_\ell \text{ occurs}\},$$

and, in agreement with Remark 2, let ν_j be the unique value of ν for which $\sigma_j = \tau_\nu$. That is, σ_j is the time at which the j^{th} regeneration level occurs. Thus, by definition,

$$S_n < \nu_j \quad \text{for } n < \tau_{\nu_j}, \qquad S_{\tau_{\nu_j}} = \nu_j$$

but

$$S_n > S_{\sigma_j} = S_{\tau_{\nu_j}} = \nu_j \quad \text{for } n > \tau_{\nu_j}.$$

Moreover, if $n = \tau_s$ but s is not one of the ν_j, then $S_t \le s$ for some $t > \tau_s$. Roughly speaking, the τ_j are the strict upward ladder epochs for the random walk $\{S_n\}$. The σ_j are special ladder epochs which make them regeneration times (in a sense to be made precise in Proposition 5). The σ_j are those ladder epochs which are visited only once. For τ_k to be such a special ladder epoch it is required that after τ_k the random walk stay strictly above its value at τ_k, that is, it is required that $\mathcal{C}_k$ occur. The special ladder epochs σ_s are regeneration epochs, because they separate the path of the random walk $\{S_n\}$ into two pieces which do not overlap (except that the endpoint of one of these pieces coincides with the initial point of the next piece).

On the event $\{\nu_j < \infty\}$ we define the j^{th} *excursion* Υ_j to be the sequence of random variables $(S_n - S_{\tau_{\nu_j}}) = (S_n - \nu_j)$, $\nu_j \le n < \nu_{j+1}$. We already proved in Remark 2 that all ν_j are finite a.s. $[Q]$. To describe the distribution of the excursions we introduce some collections of possible finite sequences which can be the value of Υ_j. For $1 \le m < \infty$, we define $\widetilde{\mathcal{M}}_m$ as the collection of sequences $\boldsymbol{\eta} = (\eta_0, \eta_1, \dots, \eta_m)$ which satisfy

$$\eta_0 = 0, \qquad \eta_i - \eta_{i-1} \in \{+1, -1\} \quad \text{for } 1 \le i \le m, \tag{118}$$

$$\text{for any } x \in \mathbb{Z}, \qquad \eta_i = x \text{ for at most } L_0 \text{ values of } i \in [0, m], \tag{119}$$

$$\eta_m > \eta_i \quad \text{for } 0 \le i < m, \tag{120}$$

but there is no $0 < j < m$ such that

$$S_i < S_j < S_\ell < S_m \quad \text{for } i < j < \ell < m. \tag{121}$$

These collections will serve to describe the distribution of Υ_j when $j = 0$. For $j \ge 1$ we shall use $\mathcal{M}_m$ which is defined as the collection of sequences $\boldsymbol{\eta} = (\eta_0, \eta_1, \dots, \eta_m) \in \widetilde{\mathcal{M}}_m$ which in addition satisfy

$$\eta_i > 0 \quad \text{for } 1 \le i \le m. \tag{122}$$

Proposition 20. *Under Q all the ν_j are a.s. finite. Moreover, the excursions Υ_j are independent, with a distribution specified by*

$$Q(\Upsilon_0 = \boldsymbol{\eta} = (\eta_0, \dots, \eta_m)) = \frac{1}{\widetilde{Z}} e^{C_4 m} P((S_0, \dots, S_m) = \boldsymbol{\eta}) \tag{123}$$

for any $\boldsymbol{\eta} \in \widetilde{\mathcal{M}}_m$. Here $\widetilde{Z}$ is a normalizing factor given by

$$\widetilde{Z} = \sum_{m=0}^{\infty} e^{C_4 m} \sum_{\boldsymbol{\eta} \in \widetilde{\mathcal{M}}_m} P((S_0, \dots, S_m) = \boldsymbol{\eta}). \tag{124}$$

Similarly, the distribution of Υ_s *with* $s \geq 1$ *is given by*

$$Q(\Upsilon_s = \boldsymbol{\eta} = (\eta_0, \dots, \eta_m)) = \frac{1}{Z} e^{C_4 m} P((S_0, \dots, S_m) = \boldsymbol{\eta}) \tag{125}$$

for any $\boldsymbol{\eta} \in \mathcal{M}_m$*, with* Z *given by*

$$Z = \sum_{m=0}^{\infty} e^{C_4 m} \sum_{\boldsymbol{\eta} \in \mathcal{M}_m} P((S_0, \dots, S_m) = \boldsymbol{\eta}). \tag{126}$$

In particular, the $(\Upsilon_s, s \geq 1)$ *are i.i.d. under* Q*. Moreover, for every* $s \geq 0$*,*

$$E^Q(\nu_{s+1} - \nu_s) < \infty. \tag{127}$$

Proof. We already know from Remark 2 that all ν_j are finite a.s. $[Q]$.

Now suppose that $\mathcal{H}(\boldsymbol{\eta})$ occurs for some $\eta \in \widetilde{\mathcal{M}}_m$. By (120) we then automatically have $m = \tau_s$ for $s = \eta_m$. Therefore, on $\mathcal{H}(\boldsymbol{\eta}) \cap \mathcal{C}_s = \mathcal{H}(\boldsymbol{\eta}) \cap \mathcal{C}_{\eta_m}$, $\tau_s \cap \mathcal{C}_s$ occurs and $s = \eta_m$ has to equal σ_r for some r and s has to be one of the ν_j. In fact, (121) shows that there can be no $j < m$ such that η_j is an earlier σ, i.e. σ_t with $t < r$. Thus, on $\mathcal{H}(\boldsymbol{\eta}) \cap \mathcal{C}_s$ we have $\sigma_1 = \tau_s$. Moreover, $\sigma_1 = \tau_s$ can occur only if $\mathcal{H}(\boldsymbol{\eta})$ for some $\boldsymbol{\eta} \in \widetilde{\mathcal{M}}_m$ occurs, as well as $\mathcal{C}_s$. Thus, $\Upsilon_0 = \boldsymbol{\eta}$ is possible only if $\boldsymbol{\eta}$ lies in $\widetilde{\mathcal{M}}_m$ for some m. Furthermore

$$\{\Upsilon_0 = \boldsymbol{\eta}\} = \{\mathcal{H}(\boldsymbol{\eta}) \cap \mathcal{C}_{\eta_m}\} \quad \text{for } \boldsymbol{\eta} \in \widetilde{\mathcal{M}}_m.$$

Also, for $\boldsymbol{\eta} \in \widetilde{\mathcal{M}}_m$, $\mathcal{H}(\boldsymbol{\eta}) \in \mathcal{G}_{\eta_m}$ (because $m = \tau_{\eta_m}$ by (120)). Hence (102), with $\mathcal{E}$ replaced by $\mathcal{H}(\boldsymbol{\eta})$, shows that

$$Q(\Upsilon_0 = \boldsymbol{\eta}) = Q(\mathcal{H}(\boldsymbol{\eta}) \cap \mathcal{C}_{\eta_m}) = C_5 e^{C_4 m} P((S_0, \dots, S_m) = \boldsymbol{\eta}, \sup_x L(m, x) \leq L_0).$$

The condition $\sup_x L(m, x) \leq L_0$ can be dropped here, because this is automatic if $S_i = \eta_i, 0 \leq i \leq m$ for some $\boldsymbol{\eta} \in \widetilde{\mathcal{M}}_m$ (by (119)). This implies (123) with (124).

To prove the statements (125) and (126) in Proposition 5 we have to show that for $\boldsymbol{\eta}^{(0)} \in \widetilde{\mathcal{M}}_{m(0)}$, and $\boldsymbol{\eta}^{(s)} \in \mathcal{M}_{m(s)}$ for $1 \leq s \leq r$, and for some constant C,

$$Q(\Upsilon_s = \boldsymbol{\eta}^{(s)}, 0 \leq s \leq r) = C \prod_{s=0}^{r} [e^{C_4 m(s)} P((S_0, \dots, S_{m(s)}) = \boldsymbol{\eta}^{(s)})]. \tag{128}$$

Let $\boldsymbol{\eta}^{(s)} = (\eta_0^{(s)} = 0, \dots, \eta_{m(s)}^{(s)})$ and write

$$q(s) = \sum_{j=0}^{s-1} m(j)$$

(with $q(0) = 0$). Then the event on the left hand side of (128) will occur if and only if

$$\mathcal{E}^{(r)} := \left\{ S_{q(s)+i} = \sum_{j=0}^{s-1} \eta_{m(j)}^{(j)} + \eta_i^{(s)},\ 0 \leq i \leq m(s), 0 \leq s \leq r \right\}, \tag{129}$$

as well as

$$\bigcap_{s=0}^{r} \mathcal{C}\Big(\sum_{j=0}^{s} \eta_{m(j)}^{(j)}\Big) \tag{130}$$

occur. Here we have written $\mathcal{C}(a)$ for $\mathcal{C}_a$ to avoid complicated subscripts. By the definitions of $\widetilde{\mathcal{M}}_m(0)$ and the $\mathcal{M}_{m(s)}$, $\eta_{m(0)}^{(0)} > 0$ and all $\eta_j^{(s)}$, $0 \le j \le m(s)$ are nonnegative. Therefore $\sum_{j=0}^{s} \eta_{m(j)}^{(j)} \ge \sum_{j=0}^{s-1} \eta_{m(j)}^{(j)}$, $1 \le s \le r$, and

$$\bigcap_{s=0}^{r-1} \mathcal{C}\Big(\sum_{j=0}^{s} \eta_{m(j)}^{(j)}\Big) \subset \mathcal{E}^{(r)} \cap \mathcal{C}\Big(\sum_{j=0}^{r} \eta_{m(j)}^{(j)}\Big).$$

Also the event (129) is contained in $\mathcal{G}(\sum_{s=0}^{r} \eta_{m(s)}^{(s)})$ (where we have written $\mathcal{G}(a)$ for $\mathcal{G}_a$), since on (129) $S_.$ reaches the level $\sum_{s=0}^{r} \eta_{m(s)}^{(s)}$ first at the time $q(r+1)$. It now follows from the fact that the value of (104) is given by (105) that the left hand side of (128) equals

$$C_5 e^{C_4 q(r+1)} P(\mathcal{E}^{(r)}, \sup_x L(q(r+1), x) \le L_0). \tag{131}$$

Finally,

$$\mathcal{E}^{(r)} = \bigcap_{s=0}^{r} \{S_{q(s)+i} - S_{q(s)} = \eta_i^{(s)},\ 0 \le i \le m(s)\},$$

and on $\mathcal{E}^{(r)}$ the range of $\{S_{q(s)+i}, 0 \le i \le m(s)\}$ consists of the integers in the interval $[\sum_{j=0}^{s-1} \eta_{m(j)}^{(j)}, \sum_{j=0}^{s} \eta_{m(j)}^{(j)}]$, $1 \le s \le r$. The interiors of these intervals are disjoint and any value x in those interiors is taken on at most L_0 times by $\{S_{q(s)+i}, 0 \le i \le m(s)\}$ if $(S_{q(s)} - S_{q(s)}, S_{q(s)+1} - S_{q(s)}, \dots, S_{q(s+1)}) - S_{q(s)}) = \boldsymbol{\eta}^{(s)}$, by virtue of (119). Moreover, on $\mathcal{E}^{(r)}$, the endpoints $\sum_{j=0}^{s} \eta_{m(j)}^{(j)}$, $0 \le s \le r$, are even taken on only once by the S_i, $0 \le i \le q(r+1)$, because $\sum_{j=0}^{s} \eta_{m(j-1)}^{(j-1)} > \sum_{j=0}^{s-1} \eta_{m(j)}^{(j)}$. Therefore, the condition $\sup_x L(q(r+1), x) \le L_0$ is automatically fulfilled on $\mathcal{E}^{(r)}$ and can be dropped from (131). The result is

$$\begin{aligned} Q(\Upsilon_r = \boldsymbol{\eta}^{(r)}, 0 \le r \le s) &= C_5 e^{C_4 q(r+1)} P(\mathcal{E}^{(r)}) \\ &= C_5 \prod_{s=0}^{r} [e^{C_4 m(s)} P(S_i = \eta_i^{(s)}, 0 \le i \le m(s))] \end{aligned} \tag{132}$$

for $\boldsymbol{\eta}^{(0)} \in \widetilde{\mathcal{M}}_{m(0)}$, $\boldsymbol{\eta}^{(s)} \in \mathcal{M}_{m(s)}$, $1 \le s \le r$. The fact that the right hand side here is a product of factors each of which depends on the value of one Υ_s only shows that the Υ_s are independent. The actual distribution of the Υ_s can also be read off from (132) and is given by (125) and (126).

Finally, the random variables $(\nu_{s+1} - \nu_s)$, $s \ge 1$, are i.i.d. under Q, so that by the renewal theorem,

$$\frac{1}{n} \sum_{\ell=1}^{n} Q(\ell \text{ equals some } \nu_j) \to [E^Q(\nu_2 - \nu_1)]^{-1} \quad \text{as } n \to \infty. \tag{133}$$

However, by (114), we know that

$$Q(\ell \text{ equals some } \nu_j) \geq Q(\mathcal{C}_\ell) \geq C_3,$$

hence

$$\liminf_{n\to\infty} \frac{1}{n} \sum_{\ell=1}^{n} Q(\ell \text{ equals some } \nu_j) \geq C_3.$$

Since $C_3 > 0$, this and (133) imply (127). □

Corollary 21.

$$E^Q(\sigma_{s+1} - \sigma_s) = E^Q(\tau_{\nu_{s+1}} - \tau_{\nu_s}) < L_0 E^Q(\nu_{s+1} - \nu_s) < \infty. \tag{134}$$

Proof. By (117), the amount of time spent by the walk in any interval $[a, b) \subset \mathbb{Z}$ is at most $L_0(b-a)$. By definition of the τ's and the σ's the walk stays in the interval $[\nu_j, \nu_{j+1})$ during $[\tau_{\nu_j}, \tau_{\nu_{j+1}}) = [\sigma_j, \sigma_{j+1})$. Thus (134) follows from (127). The strict inequality in (134) follows from the fact that for every $j \geq 0$, every site x between two successive regeneration levels $x \in [\nu_j, \nu_{j+1}) \cap \mathbb{Z}$ is visited at most L_0 times, except $x = \nu_j$ itself which is visited at most once. It follows that

$$\sigma_{j+1} - \sigma_j \leq L_0(\nu_{j+1} - \nu_j - 1) + 1$$

almost surely. Taking expectations leads to the strict inequality in (134). □

With this in mind, routine manipulations show that under Q, the position X_t satisfies the law of large numbers:

$$\lim_{n\to\infty} \frac{X_n}{n} := \gamma(L_0) \tag{135}$$

exists almost surely under Q, with

$$\gamma(L_0) := \frac{E^Q(\nu_{s+1} - \nu_s)}{E^Q(\sigma_{s+1} - \sigma_s)}.$$

By (134), we see that $\gamma(L_0) > 1/L_0$, which concludes the proof of Theorem 5. □

Acknowledgments. All our gratitude goes to Harry Kesten, whose help for Theorem 5 was invaluable. We thank A. S. Sznitman for pointing out the connection to the work on the entropic repulsion of the harmonic crystal, and Ross Pinsky for pointing out several typos in a first version of this paper. We also thank an anonymous referee for several comments which improved the paper.

References

[1] Benjamini, I., Berestycki, N.: An integral test for the transience of a Brownian path with limited local time. Ann. Inst. H. Poincaré Probab. Statist., to appear; arXiv:0806.0597v3

[2] Billingsley, P.: Convergence of Probability Measures. 2nd ed., Wiley (1999) Zbl 0944.60003 MR 1700749

[3] Bolthausen, E., Deuschel, J.-D., Giacomin, G.: Entropic repulsion and the maximum of the two dimensional harmonic crystal. Ann. Probab. **29**, 1670–1692 (2001) Zbl 1034.82018 MR 1880237

[4] Courant, R., Hilbert, D.: Methods of Mathematical Physics, Vol. I. 1st English ed., Interscience Publ. (1953) Zbl 0051.28802 MR 0065391

[5] Durrett, R.: Probability: Theory and Examples. 3rd ed., Duxbury Press MR 1609153

[6] Feller, W.: An Introduction to Probability and its Applications, Vol. I. 1st ed., Wiley (1950) Zbl 0039.13201 MR 0038583

[7] van der Hofstad, R., den Hollander, F., König, W.: Central limit theorem for the Edwards model. Ann. Probab. **25**, 573–597 (1997) Zbl 0873.60009 MR 1434119

[8] van der Hofstad, R., König, W.: A survey of one-dimensional polymers. J. Stat. Phys. **103**, 915–944 (2001) Zbl 1126.82313 MR 1851362

[9] Jost, J.: Partial Differential Equations. 2nd ed., Grad. Texts in Math. 214, Springer (2007) Zbl 1121.35001 MR 2302683

[10] Kesten, H.: On the number of self-avoiding walks. J. Math. Phys. **4**, 960–969 (1963) Zbl 0122.36502 MR 0152026

[11] Mörters, P., Sidorova, N.: A class of weakly self-avoiding walks. J. Statist. Phys. **133**, 255–269 (2008) Zbl pre05545923 MR 2448533

[12] Najnudel, J.: Construction of an Edwards' probability measure on $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$. Preprint, arXiv:0801.2751

[13] Norris, J. R., Rogers, L. C. G., Williams, D.: Self-avoiding walk: a Brownian motion model with local-time drift. Probab. Theory Related Fields **74**, 271–287 (1987) Zbl 0611.60052 MR 0871255

[14] Pinsky, R.: On the convergence of diffusion processes conditioned to remain in a bounded region for large time to limiting positive recurrent diffusion processes. Ann. Probab. **13**, 363–378 (1985) Zbl 0567.60076 MR 0781410

[15] Revuz, D., Yor, M.: Continuous Martingales and Brownian Motion. 3rd ed., Grundlehren Math. Wiss. 293, Springer (1999) Zbl 0731.60002 MR 1725357

[16] Rogers, L. C., Williams, D.: Diffusions, Markov Processes and Martingales, Vol. 2. 2nd ed., Cambridge Univ. Press (2000) Zbl 0977.60005 MR 1780932

[17] Watson, G. N.: A Treatise on the Theory of Bessel Functions. 2nd ed., Cambridge Univ. Press (1994). Zbl 0849.33001 MR 1349110