LIMIT THEOREMS FOR RANDOM WALKS ON A STRIP IN SUBDIFFUSIVE REGIMES

D. DOLGOPYAT AND I. GOLDSHEID

ABSTRACT. We study the asymptotic behaviour of occupation times of a transient random walk in a quenched random environment on a strip in a sub-diffusive regime. The asymptotic behaviour of hitting times, which is a more traditional object of study, is exactly the same. As a particular case, we solve a long standing problem of describing the asymptotic behaviour of a random walk with bounded jumps on a one-dimensional lattice. Our technique results from the development of ideas from our previous work [6] on the simple random walks in random environment and those used in [1, 2, 12] for the study of random walks on a strip.

2000 Mathematics Subject Classification: primary 60K37, 60F05; secondary 60J05, 82C44.

Keywords and Phrases: RWRE, random walks on a strip, quenched random environments, occupation times, sub-diffusive regime, products of random transformations, Lyapunov exponents.

1. INTRODUCTION

The main goal of this work is to describe the asymptotic behaviour of a random walk (RW) in a quenched random environment (RE) on a strip in a sub-diffusive regime. As a corollary we obtain a solution to a long standing problem about the asymptotic behaviour of a RW with bounded jumps in RE on a one-dimensional lattice. These two models are natural generalizations of the one-dimensional RWRE with jumps to the nearest neighbors - the so called simple RWRE (SRWRE). The techniques and ideas used in this paper resulted from the development and combination of those used in [6], where we studied the limiting behaviour of the SRWRE, and in [1, 2, 12], which studied RWRE on a strip. Our main model is the RWRE on a strip and the main quantitative characteristic of the walk that is the occupation time T_N of a large box (see (1.10) for exact definition). In [6] we also studied T_N, but on a strip the approach we use is very different from the one used for SRWRE. The important difference between SRWRE and other models can be roughly explained by the fact that a transient simple walk has to visit every point on its way to ∞, while on a strip it can miss any point with a positive probability. Due to this fact, the expectations of the occupation times of the sites form a Markov process in the 'simple' case but this is not true for a walk on a strip. In order to resolve these difficulties, we have to use methods inspired by the theory of dynamical systems such as products of random transformations, Lyapunov exponents, transfer operators combined with more probabilistic techniques such as coupling, large deviations, Poisson processes etc. We believe that the new point of view presented in this paper makes the proofs more transparent even in the classical SRWRE setting.

We now recall the exact definitions of all three models.

Model 1. In the simplest 1D case, a *random environment* is a sequence of independent identically distributed (i.i.d.) random variables $\omega = \{p_n\}_{n\in\mathbb{Z}}$, where p_n are viewed as probabilities of jumps from n to $n+1$. Given ω and $X_0 = z$, one defines a Markov chain X_t, $t = 0, 1, ...$, on $\mathbb{Z}$ with a transition kernel given by

$$\mathbb{P}_\omega(X_{t+1} = k+1|X_t = k) = p_k, \quad \mathbb{P}_\omega(X_{t+1} = k-1|X_t = k) = 1 - p_k. \tag{1.1}$$

Model 2. The RWRE on a strip $S \stackrel{\text{def}}{=} \mathbb{Z} \times \{1, \dots, m\}$ was introduced in [1] and will be the main object of our study. We say that the set $L_n \stackrel{\text{def}}{=} \{(n, j) : 1 \le j \le m\} \subset S$ is the *layer* n of the strip (or just layer n). The walker is allowed to jump from a site in L_n only to a site in L_{n-1}, L_n, or L_{n+1}. Let $X_t = (Z_t, Y_t)$ denote the coordinate of the walk at time t, where $t = 0, 1, 2, ...$, $Z_t \in \mathbb{Z}$, $1 \le Y_t \le m$.

An environment ω on a strip is a sequence of triples of $m \times m$ matrices $\omega = \{(P_n, Q_n, R_n)\}_{n\in\mathbb{Z}}$ with non-negative matrix elements and such that $P_n + Q_n + R_n$ is a stochastic matrix:

$$(P_n + Q_n + R_n)\mathbf{1} = \mathbf{1}, \tag{1.2}$$

where $\mathbf{1}$ is a vector whose all components are equal to 1. The transition kernel of the walk is given by

$$\mathbb{P}_\omega(X_{t+1} = z'|X_t = z) = \begin{cases} P_n(i,j) & \text{if } z = (n,i), z' = (n+1,j)), \\ Q_n(i,j) & \text{if } z = (n,i), z' = (n-1,j)), \\ R_n(i,j) & \text{if } z = (n,i), z' = (n,j)) \end{cases} \tag{1.3}$$

The corresponding Markov chain is completely defined if we set $X(0) = z$.

Throughout the paper we suppose that the following conditions are satisfied:

$$\{(P_n, Q_n, R_n)\}_{n\in\mathbb{Z}} \text{ is an i.i.d. sequence} \tag{1.4}$$

$$\begin{aligned} &\text{There is an } \varepsilon > 0 \text{ such that } \mathbf{P}\text{-almost surely for all } i,\, j \in [1,m] \\ &\|R_n\| < 1-\varepsilon, \quad ((I-R_n)^{-1}P_n)(i,j) > \varepsilon, \quad ((I-R_n)^{-1}Q_n)(i,j) > \varepsilon. \end{aligned} \tag{1.5}$$

Remarks. 1. The matrices P_n, Q_n, and R_n are comprised of probabilities of jumps from sites in L_n to sites in L_{n+1}, L_{n-1}, and L_n respectively. Condition (1.2) is equivalent to 'the nearest layer jumps only' property of the walk.

2. Note that $((I-R_n)^{-1}P_n)(i,j)$ and $((I-R_n)^{-1}Q_n)(i,j)$ are the probabilities for a RW starting from (n,i) to reach $(n+1,j)$ and, respectively, $(n-1,j)$ at its first exit from layer n.

3. We chose to work under conditions (1.5) in order to simplify the proofs. In fact all main results can be proved under the following much milder conditions.

$$\begin{aligned} &\text{There is } \varepsilon > 0 \text{ and integer } l \geq 1 \text{ such that } \mathbf{P}\text{-almost surely } \forall\, i \in [1,m] \\ &\left\|R_n^l\right\| < 1-\varepsilon, \quad ((I-R_n)^{-1}P_n)(i,1) > \varepsilon, \quad ((I-R_n)^{-1}Q_n)(i,1) > \varepsilon. \end{aligned} \tag{1.6}$$

Let us describe explicitly the probability spaces hidden behind the above definitions. By $(\Omega, \mathcal{F}, \mathbf{P})$ we denote the probability space describing random environments, where $\Omega = \{\omega\}$ is the set of all environments, $\mathcal{F}$ is the natural sigma-algebra of subsets of Ω and $\mathbf{P}$ is a probability measure on $(\Omega, \mathcal{F})$. The RWRE is specified by the choice of Ω and $\mathbf{P}$. Next, let $\mathfrak{X}_z = \{X(\cdot) : X(0) = z\}$ be the space of all trajectories of the walk starting from $z \in L_0$. A *quenched* (fixed) environment ω thus provides us with a conditional probability measure $\mathbb{P}_{\omega,z}$ on $\mathfrak{X}_z$ with a naturally defined probability space $(\mathfrak{X}_z, \mathcal{F}_{\mathfrak{X}_z}, \mathbb{P}_{\omega,z})$. In turn, these two measures generate a semi-direct product measure $\mathrm{P}_z := \mathbf{P} \ltimes \mathbb{P}_{\omega,z}$ which is the *annealed* probability measure on $(\Omega \times \mathfrak{X}_z, \mathcal{F} \times \mathcal{F}_{\mathfrak{X}_z})$.

The expectations with respect to $\mathbb{P}_{\omega,z}$, $\mathbf{P}$, and P_z will be denoted by $\mathbb{E}_{\omega,z}$, $\mathbf{E}$, and E_z respectively.

Remark. The notations $\mathfrak{X}_z$, $\mathbb{P}_{\omega,z}$, E_z etc. emphasize the dependence of these objects on the starting point z of the walk. However, we often use the simplified version of these notations such as $\mathbb{P}_\omega$, $\mathbb{E}_\omega$, etc. because the asymptotic behaviour of the walk does not depend on z and it is usually clear from the context what the starting point of the walk is.

Model 3. The random walk on $\mathbb{Z}$ with uniformly bounded jumps is another natural extension of the nearest neighbour model. The random environment $\omega \stackrel{\text{def}}{=} \{p(x) = (p(x,k))_{-m\leq k\leq m}\}_{x\in\mathbb{Z}}$, where $p(x)$ is a stationary in x sequence of vectors with $\sum_{k=-m}^m p(x,k) = 1$ and $p(x,k) \geq 0$. For a given environment ω the transition kernel of the walk is defined by

$$\mathbb{P}_\omega\left(X(t+1) = x+k \,|\, X(t) = x\right) = p(x,k), \quad x \in \mathbb{Z} \tag{1.7}$$

The following geometric construction transforms this walk into a walk on a strip. Let us view $\mathbb{Z}$ as a subset of the X-axis in a two-dimensional plane. Cut the X-axis into equal intervals of length m so that each of them contains exactly m consecutive integer points. Turn each such interval around its left most integer point anti-clockwise by $\pi/2$. The image of $\mathbb{Z}$ obtained in this way is a part of a strip with distances between layers equal to m. Re-scaling the X-axis of the plane by m^{-1} makes the distance between these layers equal to one and the RW on $\mathbb{Z}$ transforms into a RW on a strip with jumps to the nearest layers only. The relevant formulae for matrices $P_n,\ Q_n,\ R_n$ can be found in [1], where this construction was described in a more formal way.

It is obvious that if $p(x),\ x \in \mathbb{Z}$, is an i.i.d sequence then the just defined triples of matrices $(P_n,\ Q_n,\ R_n)$ are i.i.d. It is also easy to see that (1.5) is satisfied if for some $\varepsilon > 0$

$$\mathbf{P}\{p(x,1) > \varepsilon,\ \ p(x,-1) > \varepsilon,\ \ p(x,m) > \varepsilon,\ \ p(x,-m) > \varepsilon\} = 1. \tag{1.8}$$

A much wider class of one-dimensional RW with bounded jumps is obtained if instead of (1.8) we suppose only that

$$\mathbf{P}\{p(x,1) > \varepsilon,\ \ p(x,-1) > \varepsilon\} = 1. \tag{1.9}$$

In this case (1.5) may not be satisfied but (1.6) is satisfied.

Brief comments on the history of the subject. Two pioneering papers which initiated the development of the theory of RWRE were published in 1975 by Solomon [31] and Kesten, Kozlov, Spitzer [17]. In [31] the asymptotic properties of the SRWRE were discussed at the level of the Law of Large Numbers and the surprising fact that for a wide class of parameters the SRWRE would be escaping to ∞ at a zero speed was discovered. In [17] the *limiting distributions* of the hitting times and of the position of X were found in the *annealed* setting. The extensions of the main results from these papers to the RWRE on a strip are explained below in Theorems 3, 4, and 7.

In 1982, Sinai [29] described the asymptotic behaviour of a recurrent SRWRE. He discovered a phenomena which is now called the Sinai diffusion.

The methods used in [31, 17, 29] rely heavily on the *jumps to the nearest neighbours only* property of the walk and the limiting distributions described in [17] were obtained for *annealed* RWRE. Therefore the following questions arose and were known essentially since 1975:

1. Can one describe the limiting behaviour of the *quenched* RW at least in the case of the SRWRE (model 1)?

2. What are the analogues of (a) P-almost sure results from [31], (b) the annealed limiting statements from [17] for more general models, e. g. model 3?

3. What can be said about more general classes of environments, say stationary environments with appropriate mixing properties?

In the 1982 paper Sinai explicitly stated the questions about the possibility to extend his results to more general models, such as model 3.

The attempts to find answers to question 1 are relatively recent. We shall not discuss them here in any detail. The references concerned with SRWRE along with relevant discussion can be found in [11] and [6].

Partial answers to question 2 were obtained in [3, 4, 5, 18, 22, 21]. The discussion of these results can be found in [1, 12].

Question 3 was addressed in several publications, see e.g. [23, 1, 11, 12, 33]. And even though in [6] and in this work we consider the so called i.i.d environments (as defined above) we believe that the methods we use are useful for the analysis of RW in stationary RE satisfying appropriate mixing conditions.

Finally, let us mention several results on the RWRE on a strip which are directly related to this work. The criterion for recurrence and transience has been found in [1]. A detailed description of the limiting behaviour in the recurrent regime was given in [2]. A criterion for linear growth and the quenched (and hence annealed) Central Limit Theorem (CLT) was obtained in [12] for wide classes of environments; in particular the CLT for hitting times was established for stationary environments.

Quantities characterizing the asymptotic behaviour of a RWRE

Remember that $X_t = (Z_t, Y_t)$ is the coordinate of the walk at time t with Z_t being its $\mathbb{Z}$ component. Denote by $\tilde{T}_N$ the hitting time of layer L_N – the time at which the walk starting from a site in L_0 reaches L_N for the first time. It is both natural and in the tradition of the field to consider the understanding of the main asymptotic properties of the walk as achieved if the asymptotic behaviour of Z_t as $t \to \infty$ and $\tilde{T}_N$ as $N \to \infty$ is known.

There is of course a strong connection between the asymptotic behaviour of Z_t and $\tilde{T}_N$. Obviously $\tilde{T}_N$ is strictly monotone in N and $Z_{\tilde{T}_N} = N$. This and some other, less trivial relations between these random variables were used in a very efficient way in the study of transient RWs already in [31, 17]. In particular in [17] the asymptotic distribution of Z_t was deduced from that of $\tilde{T}_N$.

In our recent work [6] on SRWRE we studied a different quantity as the main way of describing the asymptotic behaviour of the RW. Namely, we considered the *occupation time* T_N of the interval $[0, N-1]$. The asymptotic behaviour of $\tilde{T}_N$ is exactly the same as that of T_N since $|T_N - \tilde{T}_N|$ is a stochastically bounded random variable (see Lemma 2.1 from [6]). In this paper, we study a similar quantity - the occupation time of a box $[L_0, L_{N-1}] \stackrel{\text{def}}{=} \{(n,i) : 0 \le n \le N-1\}$.

Definition. The *occupation time* T_N of the box $[L_0, L_{N-1}]$ is the total time the walk X_t starting from a site in L_0 spends on this box during its life time. In other words

$$T_N = \#\{t\colon\ 0 \le t < \infty,\ X_t \in [L_0, L_{N-1}]\}, \tag{1.10}$$

Remark. Note that $T_N \equiv T_{N,z}$ depends on the starting point z of the walk. Also, we use the convention that starting from a site z counts as one visit to z.

The paper is organized as follows. We start (Section 2) by reviewing the results from [1, 12] which are used in this paper. In Section 3 we derive formulae for the expected value of occupation times and state their asymptotic properties; the latter play a major role in the analysis of the asymptotic behaviour of the RW on a strip. In Section 4 we define traps and state the main results of the paper (Theorems 5 and 6) which are followed by Theorem 7 extending to the case of the strip the classical results from [17]. Section 5 is devoted to the proof of the properties of traps followed by the derivation of Theorem 5. The proof of Theorem 6 is given in Section 6. Since this proof is similar to that of the main result in [6], we focus our attention on the differences which are due to the fact that this time we deal with a strip. Section 8 contains the extensions of our results which are not needed in the analysis of the hitting time but are important for understanding of other properties of RWRE (cf [8, 10, 16, 20] for related work in the context of SRWRE) and will be used in the future work. The paper has four appendices containing results which are not specific to RWRE. Most of these results are not completely new, but we present them in the form convenient for our purposes. Namely, Appendix A contains the estimates of occupation times for general transient Markov chains. Standard facts about the Poisson processes and their relation to stable laws are collected in Appendix B. In Appendix C we prove a renewal theorem for a system of random contractions. The fact that the assumptions of Appendix C are applicable in our setting is verified in Section 7. Appendix D contains results about the mixing properties of random walks on the strip satisfying ellipticity conditions.

Some conventions and notations.

Letters $C,\ \bar{C},\ c, \mathbf{c}$ denote positive constants, ε is a strictly positive and small enough number, and θ is a constant from the interval $(0,1)$. The values of all these constants may be different in different sections of the paper.

$[L_a, L_b] \stackrel{\text{def}}{=} \{(n,i) : a \le n \le b\}$ is the part of the strip (a box) contained between layers L_a and L_b, where $a < b$. We use the notation $[a,b]$ and the term interval $[a,b]$ for the box $[L_a, L_b]$ when the meaning of this notation is clear from the context.

$\mathcal{F}_{a,b}$ is the σ-algebra of events depending only on the environment in $[L_a, L_b]$.

e_y is a vector whose y-th coordinate is 1 and all others are zeros.

$\mathbf{1}$ is a column vector with all components equal to 1.

If $x = (x(j))$ is a vector and $A = (a(i,j))$ a matrix we put

$$\|x\| \stackrel{\text{def}}{=} \max_j |x(j)| \ \text{ which implies } \ \|A\| = \max_i \sum_j |a(i,j)|.$$

We say that A is strictly positive (and write $A > 0$) if all its components satisfy $a(i,j) > 0$. A is called non-negative (and we write $A \geq 0$) if all $a(i,j)$ are non-negative. A similar convention applies to vectors. We shall make use of the following easy fact:

$$\text{if } \ A \geq 0 \text{ then } \ \|A\| = \|A\mathbf{1}\|\,.$$

$\mathbb{X}$ denotes the set of non-negative unit vectors, $\mathbb{X} = \{x : x \in \mathbb{R}^m,\, x \geq 0, \|x\| = 1\}$.

$E_\mu(f)$, $\nu(g)$ denote the expectations of functions f and g over measure μ and ν respectively defined on the relevant probability spaces.

We often deal with N^ε, $\ln N$, $\ln\ln N$, etc which are viewed as integer numbers. Strictly speaking, we should write $\lfloor N^\varepsilon \rfloor$, $\lfloor \ln\ln N \rfloor$, etc. However, our priority lies with the simpler notation and the exact meaning is always obvious from the context.

2. Review of related results from previous work.

The purpose of this review is to list those results from [1] and [12] which will be used in this work as well as to put the results of the present work into the right context. We note that many of the statements listed below were proved in [1, 12] under assumptions which are much milder than (1.5).

2.1. **Auxiliary sequences of matrices.** Let us fix $a \in \mathbb{Z}$ and define for $n \geq a$ two sequences of matrices: φ_n and ψ_n. To this end put $\varphi_a \stackrel{\text{def}}{=} 0$ and let ψ_a be a stochastic matrix. For $n > a$ matrices φ_n and ψ_n are defined recursively:

$$\varphi_n \stackrel{\text{def}}{=} (I - R_n - Q_n\varphi_{n-1})^{-1}P_n, \quad \psi_n \stackrel{\text{def}}{=} (I - R_n - Q_n\psi_{n-1})^{-1}P_n \tag{2.1}$$

Note that the existence of $(I - R_n - Q_n\psi_{n-1})^{-1}$ follows from (1.5).

Properties of matrices φ_n. We start with the probabilistic definition of $\varphi_n \equiv \varphi_{n,a} = (\varphi_{n,a}(i,j))$ (which implies equation (2.1) for φ_n):

$$\varphi_{n,a}(i,j) = \mathbb{P}_\omega\left(\text{RW starting from } (n,i) \text{ hits } L_{n+1} \text{ at } (n+1,j) \text{ before visiting } L_a\right).$$

Obviously these probabilities are monotone functions of a and hence the limits $\eta_n \stackrel{\text{def}}{=} \lim_{a\to-\infty} \varphi_{n,a}$ exist for all (!) environments ω. Lemma 4 in [1] implies that if (1.5) is satisfied then $\eta_n > 0$ for $\mathbf{P}$-almost every ω and for $n > a$

$$\varphi_n(i,j) > \varepsilon, \qquad \psi_n(i,j) > \varepsilon \ \text{ for } \mathbf{P}\text{-almost every } \omega. \tag{2.2}$$

Definition of matrices ζ_n. It is easy to see that since ψ_a is stochastic, so are all the ψ_n, $n > a$ (Lemma 2 in [1]). The following statement from [1] describes the $a \to -\infty$ limits of $\psi_n \equiv \psi_n(\psi_a)$ and defines a stationary sequence of stochastic matrices ζ_n.

Theorem 1. *Suppose that Condition* (1.5) *is satisfied. Then*

(a) For $\mathbf{P}$-a.e. sequence ω there exists $\zeta_n = \lim_{a\to-\infty} \psi_n(\psi_a)$, where the convergence is uniform in ψ_a and the limit ζ_n does not depend on the choice of the sequence ψ_a.

(b) The sequence $\zeta_n = \zeta_n(\omega)$, $-\infty < n < \infty$, of $m \times m$ matrices is the unique sequence of stochastic matrices which satisfies the following system of equations

$$\zeta_n = (I - Q_n\zeta_{n-1} - R_n)^{-1}P_n, \quad n \in \mathbb{Z}. \tag{2.3}$$

(c) The enlarged sequence (P_n, Q_n, R_n, ζ_n), $-\infty < n < \infty$, is stationary and ergodic.

Remark. Statements (a) and (b) imply that $\zeta_n \equiv \zeta_n(\omega)$ depends only on the "past" of the environment, namely on $\omega_{\leq n} \stackrel{\text{def}}{=} ((P_k, Q_k, R_k))_{k\leq n}$.

We need the following corollary of Theorem 1 (Remark 4 in [1]).

Corollary 2.1. *Suppose that (P, Q, R) satisfies Condition* (1.5) *(this can be any triple of matrices from the support of the distribution of (P_0, Q_0, R_0)). Then there is a unique stochastic matrix ζ such that*

$$\zeta = (I - Q\zeta - R)^{-1}P. \tag{2.4}$$

Proof. Consider the environment with transition probabilities which do not change from layer to layer and are given by matrices (P, Q, R), that is $\omega = \{(P, Q, R)\}$. Then for this single environment all conditions of Theorem 1 are satisfied. Now statement (a) implies that $\zeta_n = \zeta_{n-1}$ and setting $\zeta := \zeta_n = \zeta_{n-1}$ turns equation (2.3) into (2.4). □

The non-arithmenticity condition. We are now in a position to introduce the so called non-arithmeticity condition which will be often used in the sequel. Let (P, Q, R) and ζ be as in Corollary 2.1. Set

$$A_{(P,Q,R)} = (1 - Q\zeta - R)^{-1}Q \tag{2.5}$$

and let $e^{\lambda_{(P,Q,R)}}$ be the leading eigenvalue of $A_{(P,Q,R)}$. We say that the environment satisfies the *non-arithmeticity condition* if

$$\text{the distribution of } \lambda_{(P,Q,R)} \text{ is non-arithmetic.} \tag{2.6}$$

Vectors π_n. Our sequence of stochastic matrices ζ_n is such that $\zeta_n(i, j) \geq \varepsilon >$ for some $\varepsilon > 0$. Due to that we can always construct a sequence π_n of probability vectors such that $\pi_n = \pi_{n-1}\zeta_{n-1}$. Namely, set $\pi_{n,a} = \tilde{\pi}_a\zeta_a \ldots \zeta_{n-1}$, where $\tilde{\pi}_a$ is a probability vector.

Lemma 2.2. *If $\zeta_n(i, j) \geq \varepsilon$ for some $\varepsilon > 0$ then the following limit exists and does not depend on the choice of the sequence of probability vectors $\tilde{\pi}_a$:*

$$\pi_n \stackrel{\text{def}}{=} \lim_{a\to-\infty} \tilde{\pi}_a\zeta_a \ldots \zeta_{n-1}, \tag{2.7}$$

Moreover, for $\theta = 1 - m\varepsilon$

$$||\pi_n - \pi_{n,a}|| \leq \theta^{n-1-a} \text{ and } \pi_n(i) > \varepsilon \text{ for any } i \in [1, m]. \tag{2.8}$$

Remarks. 1. In our case vectors $\pi_n \equiv \pi(\omega_{\leq n})$ form a stationary sequence.
2. Lemma 2.2 is a well known fact which follows from the usual contracting properties of products of stochastic matrices. We state it here for future references.

Matrices A_n and Lyapunov exponents. We can finally define the following sequence of matrices:

$$A_n \stackrel{\text{def}}{=} (I - Q_n\zeta_{n-1} - R_n)^{-1}Q_n. \tag{2.9}$$

Obviously, A_n is a stationary sequence and the top Lyapunov exponent of the product of matrices A_n is defined as usual by

$$\lambda \stackrel{\text{def}}{=} \lim_{n\to\infty} \frac{1}{n} \log \|A_nA_{n-1} \ldots A_1\| . \tag{2.10}$$

It is well known (see [9]) that with $\mathbf{P}$-probability 1 the limit in (2.10) exists and does not depend on ω.

2.2. **Recurrence and transience of RWRE.** The recurrence criteria was proved in [1] for a RWRE on a strip in very general ergodic setting. We need the following particular case of this result.

Theorem 2. *Suppose that Conditions* (1.4) *and* (1.5) *are satisfied. Then the following statements hold for* $\mathbf{P}$*-a.e.* ω*,* $\mathbb{P}_\omega$*-almost surely:*
(a) $\lambda < 0$ *iff the RW is transient to the right:* $\lim_{t\to\infty} X_t = \infty$,
(b) $\lambda > 0$ *iff the RW is transient to the left:* $\lim_{t\to\infty} X_t = -\infty$,
(c) $\lambda = 0$ *iff the RW is recurrent:* $\limsup_{t\to\infty} X_t = +\infty$ *and* $\liminf_{t\to\infty} X_t = -\infty$.

Remark. The proof of Theorem 2 given in [1] contains the following useful statement: the RW is recurrent or transient to the right if and only iff

$$\lim_{a\to-\infty} \varphi_k = \zeta_k \tag{2.11}$$

2.3. **Moment Lyapunov exponents $r(\alpha)$.** From now on we consider RWRE which are transient to the right, that is $\lambda < 0$. Let us define a function whose properties are responsible for the speed of growth of our RW.

Let A_n be a sequence of matrices defined by (2.9). For $\alpha \geq 0$ put

$$r(\alpha) \stackrel{\text{def}}{=} \limsup_{n\to\infty} \left(\mathbf{E}||A_n \cdots A_1||^\alpha\right)^{\frac{1}{n}} . \tag{2.12}$$

Note that if $m = 1$ then $\zeta_n = 1$, $A_n = q_n/p_n$, and $r(\alpha) = \mathbf{E}(q_0/p_0)^\alpha$. In this form $r(\alpha)$ was first introduced in [17].

Lemma 2.3. *Suppose that* (1.5) *is satisfied. Then:*
(a) the following limit exists and is finite for every $\alpha \geq 0$:

$$r(\alpha) = \lim_{n\to\infty} \left(\mathbf{E}||A_n \cdots A_1||^\alpha\right)^{\frac{1}{n}} . \tag{2.13}$$

(b) the convergence in (2.13) is uniform in $\alpha \in [0, \alpha_0]$ for any $\alpha_0 > 0$
(c) $r'(0) = \lambda$.

2.4. **Linear and sub-linear growth of the random walk.** Let as in the Introduction $X_t = (Z_t, Y_t)$ be a random walk starting from a site $z \in L_0$, $\tilde{T}_n$ be the hitting time of layer L_n by this walk.

Theorem 3. *Suppose that* (1.5) *is satisfied and that $\lambda < 0$. Then:*
(i) $r(1) < 1$ implies that for $\mathbf{P}$-a.e. environment ω with $\mathbb{P}_{\omega,z}$-probability 1

$$\lim_{n\to\infty} n^{-1}(\tilde{T}_n - \mathbb{E}_{\omega,z}\tilde{T}_n) = 0 \quad \textit{and} \quad \lim_{n\to\infty} n^{-1}\mathbb{E}_{\omega,z}\tilde{T}_n = c > 0 \tag{2.14}$$

(ii) $r(1) \geq 1$ implies that for $\mathbf{P}$-a.e. environment ω with $\mathbb{P}_{\omega,z}$-probability 1

$$\lim_{n\to\infty} n^{-1}\tilde{T}_n = \infty. \tag{2.15}$$

The fact that $\lim_{t\to\infty} t^{-1}Z_t = c^{-1}$ when $r(1) < 1$ and that $\lim_{t\to\infty} t^{-1}Z_t = 0$ when $r(1) \geq 1$ follows from (2.14) and (2.15) respectively.

These results extend the relevant statements from [31] to the case of the strip. Further details can be found in [12].

2.5. **The diffusive regime (Central Limit Theorem) for the random walk.**

Theorem 4. *Assume that* (1.5) *is satisfied, $\lambda < 0$, and $r(2) < 1$. Then there exists D_1 such that for $\mathbf{P}$-almost every environment $\frac{\tilde{T}_n - \mathbb{E}_{\omega,z}\tilde{T}_n}{\sqrt{n}}$ converges weakly as $n \to \infty$ to a normal distribution with zero mean and variance D_1.*

Remark. It is easy to show that if in addition to the conditions of Theorem 4 also (1.4) is satisfied, then there are constants $\mathbf{c}$ and D_2 such that $\frac{\mathbb{E}_{\omega,z}\tilde{T}_n - \mathbf{c}n}{\sqrt{n}}$ converges weakly to a normal distribution with zero mean and variance D_2. Consequently $\frac{\tilde{T}_n - \mathbf{c}n}{\sqrt{n}}$ converges weakly as $n \to \infty$ in the annealed setting (that is with respect to $\mathrm{P}_z := \mathbf{P} \ltimes \mathbb{P}_{\omega,z}$) to a normal distribution with zero mean and variance $D = D_1 + D_2$.

3. Occupation times.

As stated in the Introduction, in this work the study the asymptotic behaviour of the RWRE is conducted in terms of that of the asymptotic behaviour of occupation times. In this section we derive formulae for the expectations of occupation times and discuss some of their properties. Denote the time spent by the walk at site $x = (n, y)$ by ξ_x. Obviously, the distribution of ξ_x depends on the starting point of the walk, say (k, i). Since ξ_x conditioned on the walk starting from x has a geometric distribution, it is easy to find the parameters of ξ_x also for the walk starting from arbitrary (k,i). Namely, we shall find $F_{k,i} \stackrel{\text{def}}{=} \mathbb{E}_{\omega,(k,i)}(\xi_x)$ for all (k,i) in the strip including $F_x = \mathbb{E}_{\omega,x}(\xi_x)$. Then

$$\mathbb{P}_{\omega,(k,i)}\{X \text{ reaches } x\} = \mathbb{P}_{\omega,(k,i)}\{\xi_x \geq 1\} = F_{k,i}F_x^{-1}$$

and hence

$$\mathbb{P}_{\omega,(k,i)}\{\xi_x = 0\} = 1 - F_{k,i}F_x^{-1}, \quad \mathbb{P}_{\omega,(k,i)}\{\xi_x = j\} = F_{k,i}F_x^{-2}(1-F_x^{-1})^{j-1} \quad \text{if} \quad j \geq 1.$$

The expressions for $F_{k,i}$ will be given in terms of the matrices defined in section 2. Denote by F_k the m-dimensional vector with components $F_{k,i}$, $1 \leq i \leq m$ and let $e_y \in \mathbb{R}^m$ be a vector whose y^{th} coordinate is 1 and all others are zeros.

Lemma 3.1. *Suppose that* (1.5) *holds,* $x = (n,y)$*. Then for* $\mathbf{P}$ *- almost all* ω

$$F_k \equiv F_k(n,y) = \sum_{j=n}^{\infty} \zeta_k \dots \zeta_{j-1} A_j \dots A_{n+1} u_{n,y}, \quad \textit{if } k < n \tag{3.1}$$

$$F_k \equiv F_k(n,y) = \sum_{j=k}^{\infty} \zeta_k \dots \zeta_{j-1} A_j \dots A_{n+1} u_{n,y}, \quad \textit{if } k \geq n \tag{3.2}$$

where

$$u_{n,y} = (I - Q_n\zeta_{n-1} - R_n)^{-1} e_y. \tag{3.3}$$

Remark. In the above formulae, we use the conventions that $A_j \dots A_{n+1} = I$ if $j \leq n$ and $\zeta_k \dots \zeta_{j-1} = I$ if $k \geq j$. Thus the first term in (3.1) is $\zeta_k \dots \zeta_{n-1} u_{n,y}$ and the first term in (3.2) is $A_k \dots A_{n+1} u_{n,y}$.

Proof. Consider a box $[L_a, L_b]$ with $a < k, n < b$ and a walk $\tilde{X}$ on this box starting from (k,i) with absorbtion at layers L_a and L_b. Denote by $\tilde{F}_k \stackrel{\text{def}}{=} (\tilde{F}_{k,i})_{1\leq i \leq m}$ the vector of conditional expectations, $\tilde{F}_{k,i} \stackrel{\text{def}}{=} \mathbb{E}_{\omega,(k,i)}(\xi_x^{a,b})$, where $\xi_x^{a,b}$ is the occupation time of x by the walk starting from (k,i). It is easy to see (first step analysis) that $\tilde{F}_k$ satisfy the following system of equations:

$$\tilde{F}_k = \Phi_k + P_k\tilde{F}_{k+1} + R_k\tilde{F}_k + Q_k\tilde{F}_{k-1} \text{ if } a < k < b \text{ and } \tilde{F}_a = \tilde{F}_b = 0, \tag{3.4}$$

where $\Phi_k = e_y$ if $k = n$ and $\Phi_k = 0$ otherwise. Systems of equations of this form were studied in [12]. The idea is to look for solutions to (3.4) of the form $\tilde{F}_k = \varphi_k\tilde{F}_{k+1} + d_k$ with $\varphi_a = 0$ and $d_a = 0$. A simple calculation (see Lemma 8 in [12]) shows that φ_k satisfy (2.1) and that

$$\tilde{F}_k = \varphi_k \dots \varphi_{n-1} d_n + \varphi_k \dots \varphi_n d_{n+1} + \dots + \varphi_k \dots \varphi_{b-2} d_{b-1}. \tag{3.5}$$

where

$$d_r = \tilde{u}_r + \tilde{A}_r\tilde{u}_{r-1} + \tilde{A}_r\tilde{A}_{r-1}\tilde{u}_{r-2} + \dots + \tilde{A}_r \dots \tilde{A}_j\tilde{u}_{j-1} + \dots$$

and

$$\tilde{u}_r = (I - Q_r\varphi_{r-1} - R_r)^{-1}\Phi_r, \quad \tilde{A}_r \stackrel{\text{def}}{=} (I - Q_r\varphi_{r-1} - R_r)^{-1}Q_r.$$

In our case $d_l = \tilde{A}_l \dots \tilde{A}_{n+1}\tilde{u}_n$ if $l > n$, $d_n = \tilde{u}_n$ and $d_l = 0$ otherwise which turns (3.5) into a version of (3.1), (3.2) with ζ's replaced by φ's and the sums being finite. Note that:
(a) $\lim_{a\to-\infty} \varphi_j = \zeta_j$ since $\lambda < 0$ (see (2.11) or [1]);
(b) moreover $\varphi_j \nearrow \zeta_j$ and therefore also $\tilde{A}_j \nearrow A_j$ as $a \to -\infty$;
(c) thus $\tilde{F}_k$ is monotonically increasing as a decreases or b increases with the terms of $\tilde{F}_k$ converging to the corresponding terms of F_k.
Series (3.1), (3.2) converge $\mathbf{P}$ - almost surely (once again due to $\lambda < 0$) and since they have positive entries, it follows that $\tilde{F}_k$ converges $\mathbf{P}$ - almost surely to F_k as $a \to -\infty$ and $b \to +\infty$ and this proves the Lemma. ∎

One immediate corollary from (3.2) is the estimate of the probability of return from L_{n+l} to L_n. For a $\theta_0 < 1$, set

$$\Omega_{n,l,\theta_0} = \left\{\omega : \mathbb{P}_\omega\{X \text{ visits } L_n \text{ after } L_{n+l}\} \geq \theta_0^l\right\}. \tag{3.6}$$

Lemma 3.2. *There are $C > 0$, $\theta_0, \theta_1 \in (0,1)$ such that*

$$\mathbf{P}\left\{\Omega_{n,l,\theta_0}\right\} < C\theta_1^l. \tag{3.7}$$

Proof. It follows from (3.2) that $||F_{n+l}(n,y)|| \leq \text{Const} \sum_{j=n+l}^{\infty} ||A_j \dots A_{n+1}||$. Let $\xi^{n,l}$ be the number of visits to L_n by the walk starting from L_{n+l}. Then

$$\mathbb{P}_\omega\{X \text{ visits } L_n \text{ after } L_{n+l}\} \leq \max_{z \in L_{n+l}} \mathbb{P}_{\omega,z}\{\xi^{n,l} \geq 1\} \leq \max_{z \in L_{n+l}} \mathbb{E}_{\omega,z}(\xi^{n,l})$$

$$\leq \max_y ||F_{n+l}(n,y)|| \leq \text{Const} \sum_{j=n+l}^{\infty} ||A_j \dots A_{n+1}||$$

Fix any α such that $r(\alpha) < 1$. Then for any $\theta_0 > 0$

$$\mathbf{P}\{\sum_{j=n+l}^{\infty} ||A_j \dots A_{n+1}|| \geq \theta_0^l\} \leq \theta_0^{-\alpha l} \sum_{j=n+l}^{\infty} \mathbf{E}(||A_j \dots A_{n+1}||)^\alpha \leq \text{Const}\theta_0^{-\alpha l} r(\alpha)^l.$$

and it remains to choose θ_0 so that $\theta_1 := r(\alpha)\theta_0^{-\alpha} < 1$. □

Let us now discuss a corollary which follows from (3.1). This formula will quite often be used when $n - k > cN^\varepsilon$ for some $\varepsilon > 0$ and $N \to \infty$ in which case the expression for $F_k(n,y)$ can be simplified. Namely, by Lemma 2.2

$$\zeta_k \dots \zeta_{j-1} = \boldsymbol{\pi}_j + \mathcal{O}(\theta^{j-k}), \tag{3.8}$$

where $\boldsymbol{\pi}_j$ is a rank one matrix all rows of which are equal to π_j. We can now rewrite (3.1) as

$$F_k(n,y) = \sum_{j=n}^{\infty} (\boldsymbol{\pi}_j + \mathcal{O}(\theta^{j-k}))A_j \dots A_{n+1}u_n = (1 + \mathcal{O}(\theta^{n-k})) \sum_{j=n}^{\infty} \boldsymbol{\pi}_j A_j \dots A_{n+1} u_{n,y}. \tag{3.9}$$

Fix $\varepsilon > 0$ and $\theta \in (0,1)$. Then it is easy to see that for sufficiently large N

$$\mathbf{P}\left\{\theta^{N^\varepsilon} \max_{n \in [0,N]} \sum_{j=n}^{\infty} ||A_j \dots A_{n+1}|| > N^{-50(1+s)}\right\} < N^{-100}.$$

Since errors of such order as well as events of such small probability will not play any role for our results and in our proofs (see the statements of the main theorems below), the dependence of $F_k(n,y)$ on the starting point of the walk can be neglected and this vector can be replaced by a single number

$$\rho_{n,y} := \sum_{j=n}^{\infty} \pi_j A_j \dots A_{n+1} u_n. \tag{3.10}$$

We also set

$$\rho_n := \sum_y \rho_{n,y} = \sum_{j=n}^{\infty} \pi_j A_j \dots A_{n+1} \mathbf{u}_n \tag{3.11}$$

where $\mathbf{u}_n = (I - Q_n\zeta_{n-1} - R_n)^{-1}\mathbf{1}$. Obviously, ρ_n is the expectation of the time spent by the walk in layer n.

Since A_j is a stationary sequence of positive matrices there exists a stationary sequence of vectors $v_j > 0$, $v_j \in \mathbb{X}$ and numbers $\lambda_j > 0$ such that

$$A_j v_{j-1} = \lambda_j v_j \tag{3.12}$$

and there is a sequence of functionals l_j and a $\theta < 1$ such that for any vector u

$$A_j \dots A_{n+1} u = l_n(u) \lambda_j \dots \lambda_{n+1} (1 + \mathcal{O}(\theta^{n-j})) v_j.$$

Moreover l_n is well approximated by local functions, that is, there exists $\mathcal{F}_{n,n+r}$ measurable functions $l_{n,r}$ such that $||l_n - l_{n,r}|| \leq \theta^r$ whenever $r > r_0$ (r_0 depends only on θ and ε from Condition (1.5)) .

The foregoing discussion allows us to write

$$\rho_n = l_n(\mathbf{u}_n) w_n + \mathcal{R}_n \tag{3.13}$$

where

$$w_n = \sum_{j \geq n} \lambda_j \dots \lambda_{n+1} (\pi_j, v_j) \tag{3.14}$$

and $\mathcal{R}_n$ denotes the contribution of subleading terms. Denote

$$w_{n_1, n_2} = \sum_{j=n_1}^{n_2} \lambda_j \dots \lambda_{n_1+1} (\pi_j, v_j).$$

The following lemma characterizes the tail behaviour of the distribution of w_n and thus also of $\rho_n(y)$ and ρ_n. It adjusts to our needs some well known results from [15]; the latter played a major role in many previous studies of the asymptotic behaviour of the RWRE, in particular in [17, 6]. The derivation of this lemma will be given in Section 7 using the results of Appendix C. The lemma relies on the assumption that

$$\text{there is } s > 0 \text{ such that } r(s) = 1. \tag{3.15}$$

Note that $\ln(r(\cdot))$ is a strictly convex function (see e.g. section C.3) and therefore the existence of the solution s to (3.15) implies its uniqueness. On the other hand if (3.15) has no positive solutions then $r(\alpha) < 1$ for all $\alpha > 0$. In particular, the walk is diffusive in that case in view of Theorem 4.

For the rest of the paper we suppose that (3.15) is satisfied.

Further analysis will heavily rely on the asymptotic properties of the of tails of distributions of ρ_n which will follow from those of w_n. The latter are described by the following lemma.

Lemma 3.3. *Suppose that* (1.5) *is satisfied. Then there are constants* $\bar{C}$ *and* $\bar{s} > s$ *such that*

(a) If $n_2 - n_1 > \bar{C} \ln t$ *then*

$$\mathbf{P}(w_{n_1} - w_{n_1, n_2} \geq 1 | \zeta_{n_1} = \zeta, \pi_{n_1} = \pi, v_{n_1} = v) \leq C t^{-\bar{s}}.$$

(b) If $n_2 - n_1 > \bar{C} \ln t$ *and the non-arithmeticity condition* (2.6) *holds (in addition to* (1.5)*) then there is a function* $f(\zeta, \pi, v) > 0$ *such that*

$$\mathbf{P}(w_{n_1, n_2} \geq t | \zeta_{n_1} = \zeta, \pi_{n_1} = \pi, v_{n_1} = v) \sim f(\zeta, \pi, v) t^{-s}.$$

(Here and below '$\sim$' means the bound which is uniform in the parameter involved. That is, given ε *there exists* t_0 *such that if* $t > t_0$ *and* $n_2 - n_1 > \bar{C} \ln t$ *then*

$$|t^s \mathbf{P}(w_{n_1, n_2} \geq t | \zeta_{n_1} = \zeta, \pi_{n_1} = \pi, v_{n_1} = v) - f(\zeta, \pi, v)| < \varepsilon.)$$

In particular

$$\mathbf{P}(w_n \geq t | \zeta_n = \zeta, \pi_n = \pi, v_n = v) \sim f(\zeta, \pi, v) t^{-s}.$$

(c) There exists $\bar{s} > s$ *and* $\bar{C} > 0$ *such that*

$$\mathbf{P}(\mathcal{R}_n > t | \zeta_n = \zeta, \pi_n = \pi, v_n = v) \leq C t^{-\bar{s}}.$$

The proof of Lemma 3.3 is given in Section 7.

4. Main Results.

The description of the asymptotic behaviour of T_N (defined by (1.10)) will be derived from the asymptotic properties of traps. Our first main result describes these properties. Let us introduce the exact definition of a trap on $[0, N]$.

Definition. Let $M = M_N := \ln\ln N$ and $\delta > 0$ be a given (small) number, w_n is defined by (3.14). We say that n is a *massive* site if $w_n \geq \delta N^{1/s}$. A site $n \in [0, N-1]$ is *marked* if it is massive and $w_{n+j} < \delta N^{1/s}$ for $1 \leq j \leq M$. For n marked the interval $[n-M, n]$ is called the *trap* (or $\delta N^{\frac{1}{s}}$-trap) associated to n. We call the number $\mathfrak{m}_n = \sum_{j=n-M}^{n} \rho_j$ the *mass* of the trap.

Note that this definition implies that distinct traps are disjoint.

The asymptotic distribution of traps is described by the following

Theorem 5. *Assume that the non-arithmeticity condition* (2.6) *holds. Then there exists a constant* $\mathbf{c}$ *such that the point process*

$$\left\{\left(\frac{n_j}{N}, \frac{\mathfrak{m}_{n_j}}{N^{1/s}}\right) : \ n_j \textit{ is } \delta N^{1/s}\textit{-marked and } \quad 0 \leq n_j \leq N\right\} \tag{4.1}$$

converges as $N \to \infty$ *to a Poisson process on* $[0,1] \times [\delta, \infty)$*with the measure* $\mathbf{c}s dt' \mu_\delta(dt)$*, where* μ_δ *converges to a measure with density* $\frac{\mathbf{c}s}{t^{s+1}}$ *as* $\delta \to 0$.

Remarks. 1. Each component of the point process (4.1) is itself a point process converging to a Poisson process.

2. Theorem 5 extends certain statements from [6] (see Lemma 4.4 there) to the case of the walk on a strip. It may be worth mentioning that in [6] we used the term *cluster* for what we have now decided to call a *trap*; the latter term seems to better reflect the main properties of this object.

3. The measure μ_δ will be described in more explicit terms later. However, its explicit description depends on the choice of the definition of a trap and is only important because it helps to find the limit of μ_δ as $\delta \to 0$.

The above theorem plays a major role in the description of the asymptotic behaviour of the walk in the subdiffusive regime $s \in (0, 2)$. To state our second main result, we define $\mathfrak{t}_N$ which is a normalized version of T_N, namely we set

$$\mathfrak{t}_N = \begin{cases} \frac{T_N}{N^{1/s}} & \text{if } 0 < s < 1, \\ \frac{T_N - \mathbb{E}_\omega(T_N)}{N^{1/s}} & \text{if } 1 \leq s < 2. \end{cases} \tag{4.2}$$

The definition of $\mathfrak{t}_N$ implies that complete understanding of its asymptotic properties should include the study of those of $\mathbb{E}_\omega(T_N)$. The corresponding normalized quantity is defined as follows:

$$\mathfrak{u}_N = \begin{cases} \frac{\mathbb{E}_\omega(T_N)}{N^{1/s}} & \text{if } 0 < s < 1, \\ \frac{\mathbb{E}_\omega(T_N) - u_N}{N} & \text{if } s = 1, \\ \frac{\mathbb{E}_\omega(T_N) - \mathrm{E}(T_N)}{N^{1/s}} & \text{if } 1 < s < 2, \end{cases} \tag{4.3}$$

where $u_N = N\mathbf{E}(\rho_n I_{\rho_n < x_N})$ with x_N defined by $\mathbf{P}\{\rho_n > x_N\} = N^{-1}$.

To state our next theorem, we set

$$\Theta^{(N,\delta)} = \left\{\Theta_j^{(N,\delta)} := \frac{\mathfrak{m}_{n_j}}{N^{1/s}} : 0 \leq n_j \leq N \quad \text{and } n_j \text{ is } \delta N^{1/s}\text{-marked}\right\} \tag{4.4}$$

Theorem 6. *Assume that the non-arithmeticity condition* (2.6) *holds. Then for* $0 < s < 2$ *and a* $\delta > 0$ *there is a sequence* $\Omega_{N,\delta} \subset \Omega$ *such that* $\lim_{N\to\infty} \mathbf{P}(\Omega_{N,\delta}) = 1$ *and a sequence of random point processes*

$$(\Theta^{(N,\delta)}, \Gamma^{(N,\delta)}) = \left(\left\{\Theta_j^{(N,\delta)}, \Gamma_j^{(N,\delta)}\right\}\right)$$

such that

(i) The component $\Gamma^{(N,\delta)} = \{\Gamma_j^{(N,\delta)}\}$ is a collection of asymptotically i.i.d. random variables with mean 1 exponential distribution.
(ii) The $\mathfrak{t}_N$ and $\mathfrak{u}_N$ can be presented in the following form:
(a) If $0 < s < 1$ then for $\omega \in \Omega_{N,\delta}$

$$\mathfrak{t}_N = \sum_j \Theta_j^{(N,\delta)} \Gamma_j^{(N,\delta)} + R_N, \quad \text{where} \quad R_N \geq 0 \quad \text{and} \quad \mathbf{E}(1_{\Omega_{N,\delta}} R_N) = \mathcal{O}(\delta^{1-s}), \tag{4.5}$$

$$\mathfrak{u}_N = \sum_j \Theta_j^{(N,\delta)} + \hat{R}_N, \quad \text{where } \hat{R}_N \geq 0, \ \mathbf{E}(\hat{R}_N) = \mathcal{O}(\delta^{1-s}).$$

(b) If $s = 1$ then for $\omega \in \Omega_{N,\delta}$ and a given $1/2 < \kappa < 1$

$$\mathfrak{t}_N = \sum_j \Theta_j^{(N,\delta)} (\Gamma_j^{(N,\delta)} - 1) + R_N, \quad \text{where} \quad \mathbf{E}\left[1_{\Omega_{N,\delta}} \mathbb{E}_\omega(R_N^2)\right]^\kappa = \mathcal{O}(\delta^{2\kappa-1}),$$

$$\mathfrak{u}_N = \sum_j \Theta_j^{(N,\delta)} - \bar{c}|\ln\delta| + \hat{R}_N, \quad \text{where} \quad \mathbf{E}(|\hat{R}_N|^2) = \mathcal{O}(\delta).$$

(c) If $1 < s < 2$ then for $\omega \in \Omega_{N,\delta}$

$$\mathfrak{t}_N = \sum_j \Theta_j^{(N,\delta)} (\Gamma_j^{(N,\delta)} - 1) + R_N, \quad \text{where} \quad \mathbf{E}\left[1_{\Omega_{N,\delta}} \mathbb{E}_\omega(R_N^2)\right] = \mathcal{O}(\delta^{2-s}),$$

$$\mathfrak{u}_N = \sum_j \Theta_j^{(N,\delta)} - \frac{\bar{c}}{(s-1)\delta^{s-1}} + \hat{R}_N, \quad \text{where} \quad \mathbf{E}(\hat{R}_N^2) = \mathcal{O}(\delta^{2-s}).$$

Remarks. 1. Theorem 6 was proven in [6] for SRWRE. The next two remarks are similar to those following Theorem 2 in [6]; we nevertheless believe that they are worth of being repeated.

2. The estimates of the remainders in the statements of Theorem 6 hold for all $\delta > 0$ but are not uniform in N. More precisely, e. g. the relation $\mathbf{E}(|\hat{R}_N|^2) = \mathcal{O}(\delta)$ in (b) means that for any $\delta > 0$ there is N_δ and a constant C (which does not depend on δ) such that $\mathbf{E}(|\hat{R}_N|^2) \leq C\delta$ if $N > N_\delta$.

3. The dependence of $\Theta^{(N,\delta)}$ on ω persists as $N \to \infty$ whereas $\Gamma^{(N,\delta)}$ becomes "almost" independent of ω. More precisely, for $K \gg 1$ and sufficiently large N the events $B_k^N := \{|\Theta^{(N,\delta)}| = k\}$, $0 \leq k \leq K$, form, up to a set of a small probability, a partition of Ω. If $\omega \in B_k^N$ then $\Gamma^{(N,\delta)} \equiv \Gamma^{(N,\delta)}(\omega, X)$ is a collection of k random variables which converge weakly as $N \to \infty$ to a collection of k i.i.d. standard exponential random variables. Thus the only dependence of $\Gamma^{(N,\delta)}(\omega, X)$ on ω and δ which persists as $N \to \infty$ is reflected by the fact that $|\Theta^{(N,\delta)}| = |\Gamma^{(N,\delta)}|$. (Remember that X is the trajectory of the walk and the purpose of our notation is to emphasize the dependence of $\Gamma^{(N,\delta)}$ on both ω and X.)

The following statement is a corollary of Theorem 6. In the case of the SRWRE, Theorem 7 is one of the main results of [17].

Theorem 7. *The annealed walk has the following properties:*
(a) If $s < 1$ then the distribution of $\frac{T_N}{N^{1/s}}$ converges to a stable law with index s.
(b) If $1 < s < 2$ then there is a constant u such that the distribution of $\frac{T_N - Nu}{N^{1/s}}$ converges to a stable law with index s.
(c) If $s = 1$ then there is a sequence $u_N \sim cN \ln N$ (defined as in (4.3)) such that the distribution of $\frac{T_N - u_N}{N}$ converges to a stable law with index 1.

The proof of this theorem will not be given because its derivation from Theorems 5 and 6 is easy (cf Lemma B.2 in Appendix B) and also was carried out in [6].

5. The asymptotic properties of traps and roof of Theorem 5.

5.1. **Auxiliary Lemmas.** The proof of Theorem 5 requires understanding of the asymptotic behaviour of traps. The following five lemmas describe the properties of traps that shall be used in the sequel.

We start with w_n defined by (3.14). Observe that we have

$$w_n = \lambda_n w_{n+1} + (\pi_n, v_n). \tag{5.1}$$

Lemma 5.1. *There exist $\varepsilon_1 > 0$, $\varepsilon_2 > 0$, $0 < \beta < 1$ and $C > 0$ such that for any $(\bar\zeta, \bar\pi, \bar v)$ we have*

(a) If $k \le \varepsilon_1 \ln t$ then

$$\mathbf{P}(w_n \ge t, w_{n+k} \ge t | \zeta_n = \bar\zeta, \pi_n = \bar\pi, v_n = \bar v) \le \frac{C\beta^k}{t^s};$$

(b) If $k \ge \varepsilon_1 \ln t$ then

$$\mathbf{P}(w_n \ge t, w_{n+k} \ge t | \zeta_n = \bar\zeta, \pi_n = \bar\pi, v_n = \bar v) \le Ct^{-s(\varepsilon_2+1)}.$$

Proof. For brevity we shall denote $\bar{\mathbf{P}} = \mathbf{P}(\cdot | \zeta_n = \bar\zeta, \pi_n = \bar\pi, v_n = \bar v)$.

(a) From (5.1) we have

$$w_n = \lambda_n \dots \lambda_{n+k-1} + \mathcal{O}(K^k).$$

So if $K^k \ll t$ then

$$\begin{aligned}
\bar{\mathbf{P}}(w_n \ge t, w_{n+k} \ge t) \le \quad & \bar{\mathbf{E}}(1_{\lambda_n \dots \lambda_{n+k-1} \ge 1/2} P_{\zeta_{n+k}, \pi_{n+k}, v_{n+k}}(w_{n+k} \ge t)) \\
\le \quad & \frac{\bar{\mathbf{E}}(1_{\lambda_n \dots \lambda_{n+k-1} \ge 1/2} f(\zeta_{n+k}, \pi_{n+k}, v_{n+k}))}{t^s} \\
\le \quad & \frac{C\bar{\mathbf{P}}(\lambda_n \dots \lambda_{n+k-1} \ge 1/2)}{t^s} \le \frac{C\beta^k}{t^s}.
\end{aligned}$$

(b) Consider two cases

(I) $k > \bar C \ln t$ where $\bar C$ is the constant from Lemma 3.3. Then

$$\bar{\mathbf{P}}(w_n \ge t, w_{n+k} \ge t) \le \bar{\mathbf{P}}(w_{n,n-k-1} \ge t-1, w_{n+k} \ge t) + \bar{\mathbf{P}}(w_n - w_{n,n-k-1} > 1).$$

The second term is $\mathcal{O}(t^{-\bar s})$ while the first term equals to

$$\bar{\mathbf{E}}(1_{w_{n,n+k-1} \ge t-1} \mathbf{P}_{\zeta_{n-k}, \pi_{n-k}, v_{n-k}} w_{n+k} \ge t) \le \frac{C}{t^s} \bar{\mathbf{P}}(w_{n,n+k-1} \ge t-1) \le \frac{C}{t^{2s}}.$$

(II) $\varepsilon_1 \ln t < k \le \bar C \ln t$. Fix $\tilde\varepsilon \ll 1$. Then

$$\bar{\mathbf{P}}(w_n \ge t, w_{n+k} \ge t) \le \bar{\mathbf{P}}(w_{n+k} \ge t^{1+\tilde\varepsilon}) + \bar{\mathbf{P}}(w_n \ge t, t \le w_{n+k} \le t^{1+\tilde\varepsilon}).$$

The first term is $\mathcal{O}(t^{-(1+\tilde\varepsilon)s})$ while the second term is less than

$$\bar{\mathbf{P}}(\lambda_n \dots \lambda_{n+k-1} \ge t^{-\tilde\varepsilon/2}, w_{n+k} \ge t) + \bar C \ln t \max_{1 < \le j < k} \bar{\mathbf{P}}(\lambda_n \dots \lambda_{n+j} \ge t^{1-\tilde\varepsilon}, w_{n+k} \ge t).$$

Both terms are estimated in the same way so we only discuss the first one

$$\begin{aligned}
\bar{\mathbf{P}}(\lambda_n \dots \lambda_{n+k-1} \ge t^{-\tilde\varepsilon/2}, w_{n+k} \ge t) &\le \bar{\mathbf{E}}(1_{\lambda_n \dots \lambda_{n+k-1} \ge t^{\tilde\varepsilon/2}} \mathbf{P}_{\zeta_{n-k}, \pi_{n-k}, v_{n-k}}(w_{n+k} \ge t)) \\
&\le \frac{C}{t^s} \bar{\mathbf{P}}(\lambda_n \dots \lambda_{n+k-1} \ge t^{\tilde\varepsilon/2}) \le Ct^{s(1+\varepsilon_2)}
\end{aligned}$$

as claimed. □

It may happen that not all massive sites belong to one of the clusters. This situation is controlled by the following

Lemma 5.2. *There is $\beta < 1$ such that for $n \in [0, N-1]$*

$$\mathbf{P}\left(\rho_n \ge \delta N^{1/s} \textit{ and } n \textit{ is not in a trap}\right) \le \mathrm{Const} \frac{\beta^M}{N}. \tag{5.2}$$

Proof. Suppose that n is a massive point which is not in a trap. Then consider all massive points n_i such that $n < n_1 < ... < n_k < n+M$. Note that such points exist because otherwise n would have been a marked point. Let now $n^* > n_k$ be the nearest to n_k massive point. Then by construction $n^* \geq n+M$. Also $n^* \leq n+2M$ because otherwise n_k would have been a marked point and n would belong to the n_k-trap. Hence the event

$$\{n \text{ is massive and not in a trap}\} \subset \bigcup_{n' \in [n+M, n+2M]} \{w_n \geq \delta N^{1/s},\, w_{n'} \geq \delta N^{1/s}\}.$$

By Lemma 5.1(b) we obtain

$$\mathbf{P}\left(n \text{ is massive and not in a trap}\right)$$
$$\leq \sum_{n'=n+M}^{n+2M} \mathbf{P}\left(w_n \geq \delta N^{1/s},\, w_{n'} \geq \delta N^{1/s}\right) \leq \text{Const}\frac{\beta^M}{N}$$

which proves our statement. □

Our next goal is to show that $\mathbf{P}(n \text{ is massive})$ and $\mathbf{P}(n \text{ is marked})$ are of the same order.

Lemma 5.3. *(a) For each* $l, R \geq 1$ *the following limit exists*

$$f_l(\bar\zeta, \bar\pi, \bar v, R) = \lim_{t\to\infty} t^s \mathbf{P}(w_n \geq tR, w_{n+j} < t \text{ for } j = 1 \dots l | \zeta_n = \bar\zeta, \pi_n = \bar\pi, v_n = \bar v).$$

(b) Let[1] $\bar f = \lim_{l\to\infty} f_l(\bar\zeta, \bar\pi, \bar v, R)$. *Then* $|\bar f(\bar\zeta, \bar\pi, \bar v, R) - f_l(\bar\zeta, \bar\pi, \bar v, R)| = \mathcal{O}(\theta^l)$.

Proof. By (5.1)

$$\bar{\mathbf{P}}(w_n \geq tR, w_{n+j} < t \text{ for } j = 1 \dots l) \sim$$
$$\bar{\mathbf{E}}\left(\mathbf{P}_{\zeta_{n+l}, \pi_{n+l}, v_{n+l}}\left(w_{n+l} \in \left[\frac{tR}{\lambda_n \dots \lambda_{n+l-1}}, \frac{t}{\max_j(\lambda_n \dots \lambda_{n+j-1})}\right]\right)\right)$$

so the result follows from Lemma 3.3.

(b) Since

$$\bar{\mathbf{P}}(w_n \geq Rt, w_{n+j} < t \text{ for } j = 1 \dots l) - \bar{\mathbf{P}}(w_n \geq Rt, w_{n+j} < t \text{ for } j = 1 \dots l+1)$$
$$\leq \bar{\mathbf{P}}(w_n \geq t, w_{n+l+1} \geq t)$$

the result follows by Lemma 5.1(a). □

Lemma 5.4. $\mathbf{P}(\bar f(\zeta, \pi, v, 1) > 0) > 0.$

Proof. Assume to the contrary that $\mathbf{P}(\bar f > 0) = 0$. Then for each ε there is n_0 such that for $N \geq N_0$ we have $\mathbf{P}(n \text{ is marked}) \leq \frac{\varepsilon}{N}$. Combining Lemma 3.3 and Lemma 5.2 we obtain that there is a constant c such that

$$\mathbf{P}(\mathbb{T}_n) \geq \frac{c}{N}, \text{ where } \mathbb{T}_n = \{n \text{ is marked and belongs to a trap}\}. \tag{5.3}$$

If n is in a trap let D_n be the distance to the nearest marked point to the left of n. Given D we write

$$\mathbf{P}(\mathbb{T}_n) = \mathbf{P}(\mathbb{T}_n \text{ and } D_n < D) + \mathbf{P}(\mathbb{T}_n \text{ and } D_n \geq D).$$

The first term equals to

$$\sum_{j=0}^{D-1} \mathbf{P}(n \text{ is massive and } n+j \text{ is marked}) \leq \sum_{j=0}^{D-1} \mathbf{P}(n+j \text{ is marked}) \leq \frac{\varepsilon D}{N}.$$

On the other hand

$$\mathbf{P}(\mathbb{T}_n \text{ and } D_n \geq D) \leq \sum_{j=D}^{M} \mathbf{P}(n \text{ and } n+j \text{ are massive}) \leq \frac{C\beta^D}{N}.$$

[1]the limit exists since f_l is decreasing

Thus
$$P(\mathbb{T}_n) \leq \frac{\varepsilon D + C\beta^D}{N}.$$
Taking D so large that $C\beta^D \leq \frac{c}{3}$ and then choosing $\varepsilon \leq \frac{c}{3D}$ we obtain a contradiction with (5.3). □

We now turn our attention to the mass of the trap. Observe that for $j \leq M$ we have
$$\rho_{n-j} \sim w_n \lambda_{n-j} \dots \lambda_n l_{n-j}(v_{n-j}).$$
Accordingly introduce
$$a_n = \sum_{j=0}^{\infty} \lambda_{n-j} \dots \lambda_n l_{n-j}(v_{n-j}) \text{ and } a_{n_1,n_2} = \sum_{j=0}^{n_2-n_1} \lambda_{n_2-j} \dots \lambda_{n_2} l_{n_2-j}(v_{n_2-j}).$$
In the next result proven in Section 7 we use the same notation as in Lemma 3.3.

Lemma 5.5. *(a) If $k > \bar{C} \ln t$ then*
$$\mathbf{P}(a_n - a_{n-k,n} \geq 1) \leq C t^{-\bar{s}} \text{ and}$$

(b) $\quad \mathbf{P}(a_{n-k,n} \geq t) \sim \hat{c} t^{-s}$.

(c) Moreover there exists a measure ν such that if $k > \bar{C} \ln t$ then
$$t^s \mathbf{E}\left(1_{a_{n-k,n} \geq t} f\left(\zeta_n, \pi_n, v_n, \frac{a_{n-k,n}}{t}\right)\right) \sim \hat{c} \iint f(\zeta, \pi, v, z) \frac{dz d\nu(\zeta, \pi, v)}{z^{s+1}}.$$

Corollary 5.6. *(a) The following limit exists*
$$h(t, \delta) = \lim_{N \to \infty} N^{1/s} \mathbf{P}(n \textit{ is marked and } \mathfrak{m}_n > t N^{1/s}).$$
(b) There is $\mathbf{c} > 0$ such that $\lim_{\delta \to 0} h(t, \delta) = \mathbf{c} t^{-s}$.

Proof. Take W such that if n is marked then $w_n \in [\delta N^{1/s}, \delta W N^{1/s}]$. We have
$$\mathbf{P}(n \text{ is marked and } \mathfrak{m}_n > t N^{1/s}) = \mathbf{E}\left(1_{a_{n-M,n} > \frac{t}{\delta W}} \mathbf{P}_{\zeta_n, \pi_n, v_n}\left(w_n > \min\left(1, \frac{t}{\delta a_n}\right) \delta N^{1/s}\right)\right)$$
$$\sim \frac{1}{\delta^s N} \mathbf{E}\left(1_{a_{n-M,n} > \frac{t}{\delta W}} \bar{f}\left(\zeta_n, \pi_n, v_n, \min\left(1, \frac{\delta a_n}{t}\right)\right)\right).$$
This proves (a). To prove (b) we use Lemma 5.5(c) to get
$$\frac{1}{\delta^s} \mathbf{E}\left(1_{a_{n-M,n} > \frac{t}{\delta W}} \bar{f}\left(\zeta_n, \pi_n, v_n, \min\left(1, \frac{\delta a_n}{t}\right)\right)\right)$$
$$\sim \hat{c} \frac{W^s}{t^s} \iint \bar{f}(z, \pi, v, \min(1, W/z)) \frac{d\nu dz}{z^s}.$$
□

5.2. **Proof of Theorem 5.** We are now in a position to prove Theorem 5. The following lemma essentially repeats the statement of Theorem 5 with the difference that we can now state it in terms of $h(\cdot, \delta)$ studied above (Corollary 5.6).

Lemma 5.7. *Suppose that all conditions of Theorem 5 are satisfied. Then*

(a) For a fixed $\delta > 0$ the point process
$$\{(nN^{-1}, \mathfrak{m}_n N^{-\frac{1}{s}}) : n \textit{ is } \delta N^{\frac{1}{s}}\textit{-marked}\}$$
converges as $N \to \infty$ to a Poisson process on $[0,1] \times [\delta, \infty)$ with measure $dt' \mu_\delta$ such that $\mu_\delta([t_1, t_2]) = h(t_2, \delta) - h(t_1, \delta)$.

(b) As $\delta \to 0$ μ_δ converges to a measure with density $\frac{\mathbf{c}s}{t^{s+1}}$.

Proof. To prove (a) we use Bernstein's big block-small block method. Namely, we divide $[0, N]$ into big blocks I_j of length N^{ε_3} separated by small blocks of length $N^{\varepsilon_3/2}$. We take $\varepsilon_3 < \varepsilon_2$, where ε_2 is the constant from Lemma 5.1. By Lemma 3.3 the probability that there is a trap inside the union of the short intervals tends to 0 so it suffices to consider the union of long intervals. We claim that for each j the probability that I_j contains two or more traps is $o(N^{\varepsilon_3-1})$. Indeed due to Lemma 5.1, the above probability is bounded by

$$\sum_{n\in I_j}\sum_{M_N<k<N^{\varepsilon_3}} \mathbf{P}(\text{both } n \text{ and } n+k \text{ are marked})$$

$$\leq \sum_{n\in I_j}\sum_{M_N<k<N^{\varepsilon_3}} \mathbf{P}(\text{both } n \text{ and } n+k \text{ are massive})$$

$$\leq \frac{C}{N}\sum_{n\in I_j}\left[\sum_{k=M_N}^{\varepsilon_1 \ln N}\beta^k + \sum_{k=\varepsilon_1\ln N}^{N^{\varepsilon_3}} N^{-\varepsilon_2}\right] \leq CN^{\varepsilon_3-1}\left[\beta^{M_N}+N^{\varepsilon_3-\varepsilon_2}\right]$$

proving the claim.

Next, let $\{J_l\}$ be a collection of d non-intersecting open intervals in $[\delta, \infty)$ and $K = (t_1', t_2')$ be an open interval from $[0, 1]$. Let

$$N_l = \text{Card}(n \leq N : n \text{ is marked and } (n, \mathfrak{m}_n) \in NK \times N^{1/s}J_l).$$

We need to show that, as $N \to \infty$, $\{N_l\}$ converge to independent random variables having Poisson distribution with parameter $(t_2' - t_1')\mu_\delta(J_l)$ (the proof for all other finite collections of open quadrangles in $[0, 1] \times [\delta, \infty)$ easily follows from this case). Weshall now replace $\mathfrak{m}_n$ by $\widetilde{\mathfrak{m}}_n$ which are defined similarly to $\mathfrak{m}_n$ but have the property of being i.i.d. random variables and, at the same time, $|\mathfrak{m}_n - \widetilde{\mathfrak{m}}_n| \leq N^{-\frac{100}{s}}$. Namely, we do the following:

1. Define Markov process $(\psi_n, \hat{\pi}_n)$ starting with initial conditions $\hat{\pi}_{n_j}$ and ψ_{n_j} which have all entries equal to $\frac{1}{m}$, where n_j is the middle of the short interval preceding I_j. This process is defined for $n > n_j$ with the ψ component given by (2.1) and $\hat{\pi}_n = \hat{\pi}_{n-1}\psi_{n-1}$.
2. Similarly to (3.11), set for $n \in I_j$

$$\hat{\rho}_n := \sum_{j=n}^{n_{j+1}} \hat{\pi}_j \hat{A}_j \dots \hat{A}_{n+1}\hat{\mathbf{u}}_n,$$

where $\hat{A}_j$, $\hat{\mathbf{u}}_n$ are define as the corresponding As and $\hat{\mathbf{u}}$s with ζs replaced by ψs.

3. Finally $\widetilde{\mathfrak{m}}_n$ is defined similarly to $\mathfrak{m}_n$ but with $\hat{\rho}_n$ replacing ρ_n.

The independence mentioned above is obvious from the construction and the approximation property now follows from Proposition D.1:

$$||\zeta_n - \psi_n|| \leq \text{Const}\theta^{n-n_j}.$$

Let now $\{\Gamma_j\}$ be a sequence of random d-dimensional vectors such that

$$\Gamma_j = \begin{cases} e_l & if I_j \subset NK, \text{ there is exactly one trap } \tilde{n} \in I_j, \ \mathfrak{m}_{\tilde{n}} \in N^{\frac{1}{s}}J_l, \\ 0 & \text{in all other cases.} \end{cases}$$

Then Γ_j are iid random vectors and

$$\mathbf{P}(\Gamma_l = e_l) \sim (t_2' - t_1')\mu_\delta(J_l)N^{\varepsilon_3-1}, \quad \mathbf{P}(\Gamma_l = 0) = 1 - (t_2' - t_1')\mu_\delta(J_l)N^{\varepsilon_3-1} + o(N^{\varepsilon_3-1}).$$

Therefore part (a) follows from the Poisson Limit Theorem for independent random vectors. Part (b) follows from part (a) and Corollary 5.6(b). □

6. Proof of Theorem 6.

The proof of Theorem 6 is to an extent similar to that of Theorem 2 in [6] and we shall give only an outline of it putting emphases at those parts which are new. As in [6], we start with a lemma which allows us to show the smallness of the contribution to T_N which comes from the sites where $\rho_n < \delta N^{\frac{1}{s}}$. We then compute the main contribution to T_N which comes from the traps in $[0, N]$ described in Theorem 5.

Within this section we shall use the following notation: $\xi_n = \sum_{i=1}^m \xi_{(n,i)}$. Obviously, $\rho_n = \mathbb{E}_\omega(\xi_n)$. In these notations, Lemma 6.1 becomes an exact copy of Lemma 4.1 from [6].

Lemma 6.1. *Let $\delta > 0$. Then there is N_δ such that for $N > N_\delta$ the following holds:*

(a) If $0 < s < 1$ then

$$\mathrm{E}\left(\sum_{w_n < \delta N^{1/s}} \xi_n\right) \leq \mathrm{Const} N^{1/s} \delta^{1-s}.$$

(b) If $1 < s < 2$ then there is a set $\tilde{\Omega}_{N,\delta}$ such that $\mathbf{P}(\tilde{\Omega}^c_{N,\delta}) \leq N^{-100}$ and

$$\mathbf{E}\left(1_{\tilde{\Omega}_{N,\delta}} \mathbb{E}_\omega \left(\sum_{w_n < \delta N^{1/s}} (\xi_n - \rho_n)\right)^2\right) \leq \mathrm{Const} N^{2/s} \delta^{2-s}.$$

(c) If $0 < s < 1$ then

$$\mathbf{E}\left(\sum_{w_n < \delta N^{1/s}} \rho_n\right) \leq \mathrm{Const} N^{1/s} \delta^{1-s}.$$

(d) If $1 < s < 2$ then

$$\mathrm{Var}\left(\sum_{w_n < \delta N^{1/s}} \rho_n\right) \leq \mathrm{Const} N^{2/s} \delta^{2-s}.$$

(e) If $s = 1$ then given $\frac{1}{2} < \kappa < 1$ there is a set $\tilde{\Omega}_{N,\delta}$ such that $\mathbf{P}(\tilde{\Omega}^c_{N,\delta}) \leq N^{-100}$ and

$$\mathbf{E}\left(1_{\tilde{\Omega}_{N,\delta}} \left(\mathrm{Var}_\omega \left(\sum_{w_n < \delta N} (\xi_n - \rho_n)\right)\right)^\kappa\right) \leq \mathrm{Const} N^{2\kappa} \delta^{2\kappa - 1}, \tag{6.1}$$

$$\mathbf{E}\left(\left(\sum_{w_n < \delta N} (\rho_n - \mathbf{E}(\rho I_{\rho < \delta N}))\right)^2\right) \leq \mathrm{Const} N^2 \delta. \tag{6.2}$$

Proof. Parts (a) and (c) follow from Lemma 3.3 and Markov inequality (cf. the proof of Lemma 4.1 in [6]).

The proofs of (b), (d), and (e) in Lemma 4.1 in [6] do not go through directly in the case of the strip. We shall give a complete proof of (b). The required adaptations in the cases (d) and (e) are the same.

Proof of (b). Set $\chi_n = I_{w_n < \delta N^{1/s}}$; this notation will be used only within the proof of Lemma 6.1. Denote $\tilde{Y}_\delta = \sum_{w_n < \delta N^{1/s}} (\xi_n - \rho_n)$. Then $\mathbb{E}_\omega(\tilde{Y}_n) = 0$ and so it suffices to show that $\mathrm{Var}_\omega(\tilde{Y}_\delta) = \mathcal{O}(\delta^{2-s} N^{2/s})$ except for ω from a set of small probability. It follows from Lemma 3.2 that if K is sufficiently large, $n_2 - n_1 > K \ln N$, and $\omega \notin \Omega_{n_1, K \ln N, \theta_0}$ (see (3.6) for the definition of this set) then

$$\mathrm{Cov}_\omega(\xi_{n_1}, \xi_{n_2}) \leq C N^{-100}.$$

Therefore

$$\begin{aligned}\mathrm{Var}_\omega(\tilde{Y}_\delta) &= \left| o(1) + \sum_{n_2 - K\ln N < n_1 < n_2} 2\chi_{n_1}\chi_{n_2}\mathrm{Cov}_\omega\left(\xi_{n_1}, \xi_{n_2}\right) + \sum_n \chi_n \mathrm{Var}_\omega\left(\xi_n\right)\right| \\ &\le 1 + \mathrm{Const} \sum_{n_2 - K\ln N < n_1 \le n_2} \rho_{n_1}\rho_{n_2}\chi_{n_1}\chi_{n_2}\end{aligned} \tag{6.3}$$

where the summation is over pairs with $\rho_{n_i} < \delta N^{1/s}$. The last step uses

$$|\mathrm{Cov}_\omega\left(\xi_{n_1}, \xi_{n_2}\right)| \le \sqrt{\mathrm{Var}_\omega\left(\xi_{n_1}\right)\mathrm{Var}_\omega\left(\xi_{n_2}\right)} \le C\rho_{n_1}\rho_{n_2}.$$

Here, apart of the Cauchy-Schwartz inequality, we use the fact that if the walk starts from $z \in L_{n_0}$, then $\mathbb{P}_{\omega,z}\{\xi_{n,j} \ge 1\} \ge \varepsilon_0$ for $n \ge n_0$ (the existence of ε_0 is due to (1.5)). The latter inequality implies that $\sqrt{\mathrm{Var}_\omega\left(\xi_{n,i}\right))} \le \mathrm{Const}\rho_{n,i}$ with the constant depending only on ε_0. We use here elementary explicit expressions for all involved quantities, see (A.8) and (A.15).

Next, we have to estimate the expectation of the last sum in (6.3). To this end introduce

$$\mathfrak{r}_n = 1 + ||A_{n+1}|| + ||A_{n+2}A_{n+1}|| + ... \equiv \sum_{j=0}^{\infty} ||A_{n+j}...A_{n+1}||.$$

It is clear from (3.11) and the strong ellipticity condition (1.5) that there are constants $c_1,\ c_2$ such that $c_1\mathfrak{r}_n < w_n < c_2\mathfrak{r}_n$. Hence there is a C and t_0 such that for $t > t_0$ uniformly in ζ

$$\mathbf{P}\{\mathfrak{r}_n > t|\zeta_n = \zeta\} \le Ct^{-s} \text{ and } \mathbf{E}\left[\left(\mathfrak{r}_n^2|\zeta_n = \zeta\right) I_{(\mathfrak{r}_n|\zeta_n=\zeta)\le t}\right] \le \mathrm{Const}\, t^{2-s}. \tag{6.4}$$

We also have that for $k > 0$

$$\rho_{n-k} \le C\sum_{j=0}^{\infty} ||A_{n-k+j}...A_{n-k+1}|| \le C\sum_{j=0}^{k-1} ||A_{n-k+j}...A_{n-k+1}|| + C||A_n...A_{n-k+1}||\mathfrak{r}_n.$$

and

$$\rho_{n-k}\rho_n \le C\sum_{j=0}^{k-1} ||A_{n-k+j}...A_{n-k+1}||\mathfrak{r}_n + C||A_n...A_{n-k+1}||\mathfrak{r}_n^2.$$

To estimate $\mathbf{E}(\mathfrak{r}_{n-k}\mathfrak{r}_n)$ we condition on ζ_n and use the fact that the conditioned random variables $||A_{n-k+j}...A_{n-k+1}|||\zeta_n$ and $\mathfrak{r}_n|\zeta_n$ are independent. Therefore

$$\mathbf{E}\left(||A_{n-k+j}...A_{n-k+1}||\mathfrak{r}_n\right) = \mathbf{E}\left[\mathbf{E}\left(||A_{n-k+j}...A_{n-k+1}|||\zeta_n\right)\mathbf{E}\left(\mathfrak{r}_n|\zeta_n\right)\right] \le C\beta^j,$$

where $\beta := r(1) < 1$ and $\mathbf{E}\left(\mathfrak{r}_n|\zeta_n\right) < \mathrm{Const}$ since $s > 1$.

Similarly, but this time using also (6.4) we obtain

$$\begin{aligned}\mathbf{E}\left(||A_n...A_{n-k+1}||\mathfrak{r}_n^2\chi_n\right) &\le \mathbf{E}\left[\mathbf{E}\left(||A_n...A_{n-k+1}|||\zeta_n\right)\mathbf{E}\left(\mathfrak{r}_n^2 I_{(\mathfrak{r}_n|\zeta_n=\zeta)\le C\delta N^{\frac{1}{s}}}|\zeta_n\right)\right] \\ &\le C\beta^k\delta^{2-s}N^{\frac{2-s}{s}}.\end{aligned}$$

Hence

$$\mathbf{E}\left(\rho_{n-k}\rho_n\chi_n\right) \le \mathrm{Const}\left(1 + \beta^k\delta^{2-s}N^{\frac{2-s}{s}}\right)$$

and therefore for $N > N_\delta$

$$\mathbf{E}\left(\sum_{k=0}^{K\ln N}\left(\rho_{n-k}\rho_n\chi_n\right)\right) \le \mathrm{Const}\left(\ln N + \delta^{2-s}N^{\frac{2-s}{s}}\right) \le C\delta^{2-s}N^{\frac{2-s}{s}}.$$

Finally

$$\begin{aligned}\mathbf{E}\left(\sum_{n_2-K\ln N<n_1\le n_2}\rho_{n_1}\rho_{n_2}\chi_{n_1}\chi_{n_2}\right) &\le \mathbf{E}\left(\sum_{n_2-K\ln N<n_1\le n_2}\rho_{n_1}\rho_{n_2}\chi_{n_2}\right) \\ &\le CN\delta^{2-s}N^{\frac{2-s}{s}} = C\delta^{2-s}N^{2/s}.\end{aligned}$$

Part (b) is proven. The proofs of parts (d) and (e) follow the proof of the corresponding statement of Lemma 4.1 in [6] with the modifications similar to ones presented here. □

We are now prepared to explain the main steps of the proof of Theorem 6.

We consider the case $s \in (0,1)$; other cases are treated similarly. Present the time spent by the walk in $[L_0, L_{N-1}]$ as

$$T_N = \sum_{n=0}^{N-1} \xi_n \equiv \sum_{n=0}^{N-1}\sum_{i=1}^{m} \xi_{(n,i)} = S_1 + \ S_2 + \ S_3, \tag{6.5}$$

where

$$S_1 = \sum_{n:\, w_n<\delta N^{1/s},\, n\notin \text{any trap}} \xi_n$$

$$S_2 = \sum_{n:\, w_n\geq\delta N^{1/s},\, n \text{ is not in a trap}} \xi_n$$

$$S_3 = \sum_{n:\, n \text{ is in a trap}} \xi_n .$$

By Lemma 6.1, (a) we have that $\mathbf{E}(S_1) \leq \mathrm{Const} N^{1/s}\delta^{1-s}$.

Next, denote

$$\bar\Omega^{(1)}_{N,\delta} := \{\omega : \text{ there is } n \in [0, N-1] \text{ s. t. } w_n \geq \delta N^{1/s},\ n \text{ is not in a trap}\}.$$

It follows from (5.2) that

$$\mathbf{P}\left(\bar\Omega^{(1)}_{N,\delta}\right) \leq N\mathbf{P}\left(w_n \geq \delta N^{1/s} \text{ and } n \text{ is not in a trap}\right) \leq \mathrm{Const}\beta^M.$$

But then

$$\mathbf{P}(S_2 \neq 0) \leq \mathbf{P}\left(\bar\Omega^{(1)}_{N,\delta}\right) \leq \mathrm{Const}\beta^M \to 0 \text{ as } N\to\infty.$$

We thus have that for $\omega \notin \bar\Omega^{(1)}_{N,\delta}$

$$\mathbf{t}_N = N^{-\frac1s} S_3 + N^{-\frac1s} S_1 = N^{-\frac1s} S_3 + R_N,$$

where $R_N := N^{-\frac1s} S_1$ and satisfies the requirements of (a), Theorem 6.

It remains to analyze S_3 which is the main contributor to T_N coming from the sum over the traps in $[L_0, L_{N-1}]$. Let us present it as follows:

$$N^{-\frac1s} S_3 = \sum_{n:\, n \text{ is marked}} N^{-\frac1s} \sum_{j=0}^{M} \xi_{n-j}.$$

Since n is marked, we can choose a k such that $\rho_{n,k} \geq m^{-1}\delta N^{\frac1s}$. Now present

$$\sum_{j=0}^{M} \xi_{n-j} = \sum_{j=0}^{M}\sum_{i=1}^{m}\left(\frac{\xi_{n-j,i}}{\rho_{n-j,i}} - \frac{\xi_{n,k}}{\rho_{n,k}}\right)\rho_{n-j,i} + \frac{\xi_{n,k}}{\rho_{n,k}}\sum_{j=0}^{M}\rho_{n-j}.$$

Next, we shall use Corollary A.2 to estimate $\left\|\frac{\xi_{n-j,i}}{\rho_{n-j,i}} - \frac{\xi_{n,k}}{\rho_{n,k}}\right\|$, where $\|f\| := \sqrt{\mathbb{E}_\omega(|f|^2)}$ for a function f on the space of trajectories of the walk. We have:

$$\left\|\frac{\xi_{n-j,i}}{\rho_{n-j,i}} - \frac{\xi_{n,k}}{\rho_{n,k}}\right\| \leq \sum_{r=n-j}^{n-1}\left\|\frac{\xi_{r,i}}{\rho_{r,i}} - \frac{\xi_{r+1,i}}{\rho_{r+1,i}}\right\| + \left\|\frac{\xi_{n,i}}{\rho_{n,i}} - \frac{\xi_{n,k}}{\rho_{n,k}}\right\| \leq C\sum_{r=n-j}^{n-1}\frac{1}{\sqrt{\rho_{r,i}}} + \frac{C}{\sqrt{\rho_{n,k}}}.$$

Condition (1.5) together with (3.1) imply that there is $\varepsilon_0 > 0$ such that $\rho_{n-j,i} \geq \varepsilon_0 ||A_{n-j} \dots A_n|| \rho_{n,k}$. Hence for $n-j$ belonging to a trap, that is $(n-j) \in [n-M, n]$, we have that $\rho_{n-j,i} \geq c\varepsilon_0^M \rho_{n,k} \geq cN^{-\bar\varepsilon}\rho_n$. (Remember that $M = \ln\ln N$ and therefore these inequalities hold for any $\bar\varepsilon > 0$ and $N > N_{\bar\varepsilon}$.) Thus

$$\left\| \sum_{j=0}^{M} \sum_{i=1}^{m} \left(\frac{\xi_{n-j,i}}{\rho_{n-j,i}} - \frac{\xi_{n,k}}{\rho_{n,k}} \right) \rho_{n-j,i} \right\| \leq \text{Const} \frac{N^{\bar\varepsilon/2}}{\sqrt{\rho_n}} \sum_{j=1}^{M} M\rho_{n-j} \leq \text{Const} \frac{N^{\bar\varepsilon}}{\sqrt{\rho_n}} \sum_{j=1}^{M} \rho_{n-j}$$

If for n marked we set

$$\mathfrak{a}_n = \mathfrak{m}_n^{-1} \sum_{j=0}^{M} \sum_{i=1}^{m} \left(\frac{\xi_{n-j,i}}{\rho_{n-j,i}} - \frac{\xi_{n,k}}{\rho_{n,k}} \right) \rho_{n-j,i}$$

then $\|\mathfrak{a}_n\| \leq \text{Const}\frac{N^{\bar\varepsilon}}{\sqrt{\rho_n}} \to 0$ as $N \to \infty$ and we have

$$\frac{\sum_{j=0}^{M} \xi_{n-j}}{N^{1/s}} = \left(\frac{\xi_{n,k}}{\rho_{n,k}} + \mathfrak{a}_n \right) \frac{\mathfrak{m}_n}{N^{1/s}}.$$

Set $\Gamma_j^{(N,\delta)} = \frac{\xi_{n_j,k_j}}{\rho_{n_j,k_j}} + \mathfrak{a}_{n_j}$, where $\{n_j\}$ is the collection of marked points. To finish the proof of statement (a) from Theorem 6, it remains to check that

$$\{\xi_{n_j,k_j}/\rho_{n_j,k_j}\}_{n_j \text{ is marked}}$$

form a collection of asymptotically independent random variables which also are asymptotically exponential with mean 1.

The convergence to the exponential distribution is an immediate corollary of two facts: the conditional random variable $\xi_{n,k}|(\xi_{n,k} \geq 1)$ is geometric and $\mathbb{P}_\omega(\xi_{n,k} = 0) \to 0$ as $N \to \infty$ (to prove this last assertion apply (A.19) with $a = (n,k)$ and b the first point visited by the walker inside layer L_n).

To establish the asymptotic independence remember the construction used in the proof of Theorem 5. We have established there that the marked points belong to the blocks of length N^{ε_3} which are separated from each other by the blocks of length $N^{\frac{\varepsilon_3}{2}}$ and, moreover, there is at most one marked point in a large block. By Lemma 3.2, the $\mathbb{P}_\omega$ probability that the walk would ever return to block I_{j-1} after having reached I_j is of order $\mathcal{O}\left(\theta_0^{N^{\frac{\varepsilon_3}{2}}}\right)$, where $\theta_0 < 1$, it follows that any random variables which are functions of the part of trajectory of the walk starting at the left end of I_j and restricted to the $N^{\varepsilon_3/2}/2$ neighbourhood of I_j are independent.

Part (a) of Theorem 6 is proved. Parts (b) and (c) are dealt with in a similar way.

7. Tail asymptotics.

Proof of Lemma 3.3. Parts (a) and (b) of the lemma follow from the z^+ part of Theorem 9 in appendix C applied to the following Markov process:

$$\Phi_n = (\pi_n, v_n, \zeta_n), \quad g_n = (P_{n+1}, Q_{n+1}, R_{n+1}). \tag{7.1}$$

Note that due to (1.5) there exists $\bar\varepsilon$ such that both A_n and ζ_n map $\mathbb{X}$ into $\mathbb{X}_{\bar\varepsilon} = \{v \in \mathbb{X} : v_j \geq \bar\varepsilon\}$. In (7.1) π_n and v_n are regarded as elements of $\mathbb{X}_{\bar\varepsilon}$ and ζ_n is an element of the set of stochastic matrices. Recall that given Φ_n and g_n we can reconstruct Φ_{n+1} using (2.3), (2.7), (2.9) and (3.12).

In order to apply Theorem 9 we need to check three conditions. The first one is eventual contraction (equation (C.1)). Since both A_n and ζ_n map $\mathbb{X}$ into $\mathbb{X}_{\bar\varepsilon}$ we can apply Birkhoff Theorem (see e.g. [22]) which tells us that there is a constant $\bar\theta = \bar\theta(\bar\varepsilon) < 1$ such that A_n and ζ_n contract the Hilbert metric on $\mathbb{X}_{\bar\varepsilon}$ at least by factor $\bar\theta$ (the contraction of π_n part also follows from Lemma 2.2). The contraction of ζ_n part is proven in Appendix D.

Second, we need to check (C.2). In our setting we have to show that for each t there is n such that $\mathbf{P}(||A_n \dots A_1|| > t) > 0$. If this were false then there would exist $t_0 > 1$ such that $||A_n \dots A_1||^\alpha < t_0^\alpha$ for all n with probability 1. This would imply $r(\alpha) \leq 1$ for all $\alpha > 0$. Since $r(0) = 1$ and $\ln r(\alpha)$ is strictly convex we would actually have $r(\alpha) < 1$ for all positive α contradicting (3.15).

Lastly we need to show that (C.3) has no solutions. In our setting (C.3) takes form

$$e^{i\bar{u}}\mathfrak{h}(\pi_{n-1}, v_{n-1}, \zeta_{n-1}) = e^{iu\ln||A_n v_{n-1}||}\mathfrak{h}(\pi_n, v_n, \zeta_n). \tag{7.2}$$

Take (P, Q, R) in the support of the environment measure. Let $\zeta_{(P,Q,R)}$ and $A_{(P,Q,R)}$ be defined by (2.4) and (2.5) respectively and denote by $\pi_{(P,Q,R)}$ and $v_{(P,Q,R)}$ the positive eigenvectors of these matrices. Then

$$\Phi_n \equiv (\pi_{(P,Q,R)}, v_{(P,Q,R)}, \zeta_{(P,Q,R)}), \quad g_n \equiv (P, R, R)$$

is an admissible trajectory. Evaluating (7.2) along this trajectory we get

$$e^{i\bar{u}}\mathfrak{h}(\pi_{(P,Q,R)}, v_{(P,Q,R)}, \zeta_{(P,Q,R)}) = e^{iu\ln\lambda_{(P,Q,R)}}\mathfrak{h}(\pi_{(P,Q,R)}, v_{(P,Q,R)}, \zeta_{(P,Q,R)}).$$

From this we conclude that

$$\lambda_{(P,Q,R)} - \frac{\bar{u}}{u} \in \frac{2\pi}{u}\mathbb{Z}$$

contradicting the non-arithmeticity condition (2.6).

Hence Theorem 9 is applicable giving parts (a) and (b) of the lemma. To prove part (c) note that

$$\mathcal{R}_n = \sum_{j\geq n} \pi_j A_j \dots A_{n+1}[\mathbf{u}_n - l_n(\mathbf{u}_n)v_n].$$

Pick a small $\tilde{\varepsilon}$ and split $\mathcal{R}_n = \mathcal{R}'_n + \mathcal{R}''_n$ where the first term contain the terms with $j < n + \tilde{\varepsilon}\ln t$ and the second term contain the terms with $j \geq n + \tilde{\varepsilon}\ln t$. Choosing $\tilde{\varepsilon}$ small enough we can ensure that $\mathcal{R}'_n \leq \frac{t}{2}$. On the other hand for terms in $\mathcal{R}''_n$ we have $\theta^{n-j} \leq t^{-\tilde{\varepsilon}|\ln\theta|}$ and hence $\mathcal{R}''_n < Ct^{-\tilde{\varepsilon}|\ln\theta|}w_n$. Thus

$$\mathbf{P}(\mathcal{R}_n > t) \leq \mathbf{P}\left(w_n > Ct^{1+\tilde{\varepsilon}|\ln\theta|}\right)$$

and so part (c) follows from part (b). □

Proof of Lemma 5.5. The result follows from z^- part of Theorem 9 applied to the same Markov process (7.1) as in the proof of Lemma 3.3. □

8. Extensions.

Here we discuss some extensions of our results which are not used in the proof of Theorem 6 but are helpful in studying other properties of the walk. Applications of these results will be presented in a separate paper.

8.1. **Environment inside the trap.** Fix $R \geq 1$. Let $\mathbb{T}_j = [n_j - M_N, n_j]$ be the j-th trap. We call $\bar{n}_j \in \mathbb{T}_j$ *R-center of* $\mathbb{T}_j$ if $\bar{n}_j$ is the rightmost point in $\mathbb{T}_j$ such that $w_{\bar{n}_j} > w_n/R$ for all $n \in \mathbb{T}_j$. We choose R so that for each k we have

$$\mathbf{P}(\lambda_n \dots \lambda_{n+k} \in \{R, R^{-1}\}) = 0.$$

In particular, if for each k we have $\mathbf{P}(\lambda_n \dots \lambda_{n+k} = 1) = 0$ then we can take $R = 1$ so that $\bar{n}_j$ will be the point with the maximal value of w_n. Denote $\omega^{(j)} = \tau^{\bar{n}_j}\omega$, where τ is the standard shift on the space of environments. Theorem 5 can be strengthened in the following way.

Theorem 8. *Assume that the non-arithmeticity condition* (2.6) *holds. Then there exists a probability measure $\tilde{\nu}_\delta$ on Ω and a constant $\mathbf{c}$ such that the point process*

$$\left\{\left(\frac{\bar{n}_j}{N}, \frac{\mathfrak{m}_{n_j}}{N^{1/s}}, \omega^{(j)}\right)\right\} \tag{8.1}$$

converges as $N \to \infty$ to a Poisson process on $[0, 1] \times [\delta, \infty) \times \Omega$ with the measure $\mathbf{c}dt'd\mu_\delta(\mathfrak{m})d\nu_\delta(\omega)$. As $\delta \to 0$ μ_δ converges to a measure with density $\frac{\mathbf{c}s}{\mathfrak{m}^{s+1}}$ and $\tilde{\nu}_\delta$ converges to some measure $\tilde{\nu}$.

In other words if the walker is trapped, then he sees the environment distributed according to a measure $\tilde{\nu}$. This statement extends the results obtained in [13, 16].

To prove Theorem 8 we first show that as $N \to \infty$ and then $\delta \to 0$

$$\left\{\left(\frac{\bar{n}_j}{N}, \frac{w_{\bar{n}_j}}{N^{1/s}}, \omega^{(j)}\right)\right\} \text{ converges to a Poisson process with measure } \tilde{c}dt'\frac{dw}{w^{1+s}}d\tilde{\nu}^*. \tag{8.2}$$

The proof of this result is similar to the proof of Theorem 5. Namely, call n (R, l)*–maximal* if $w_n > w_{n+k}/R$ for $0 < k \le l$ and for each $0 < k' < l$ there exists $|k''| \le l$ such that $w_{n+k'} \le w_{n+k''}/R$.

Using the same argument as in Lemma 5.3 one can show that

$$\mathbf{P}\left(\frac{w_n}{t} \in [\bar{a}, \bar{\bar{a}}] \text{ and } n \text{ is } (R,l) - \text{maximal} | \zeta_n = \bar{\zeta}, \pi_n = \bar{\pi}, v_n = \bar{v}\right) \sim \hat{f}_l(\bar{\zeta}, \bar{\pi}, \bar{v}) \left(\bar{a}^{-s} - \bar{\bar{a}}^{-s}\right) t^{-s}$$

and $\hat{f}_l \to \hat{f}$ as $l \to \infty$. Moreover similarly to Lemma 5.4 one can show that $\mathbf{P}(\hat{f} > 0) > 0$ (otherwise we would get a contradiction with the fact that each trap has a center). In addition we have that for each $\bar{g}_{-l} \dots \bar{g}_l$

$$\mathbf{P}\left(\frac{w_n}{t} \in [\bar{a}, \bar{\bar{a}}] \text{ and } n \text{ is } (R,l) - \text{maximal} | \zeta_n = \bar{\zeta}, \pi_n = \bar{\pi}, v_n = \bar{v}, g_{n-l} = \bar{g}_{-l}, \dots g_{n+l} = \bar{g}_l\right)$$

$$\sim \tilde{f}(\bar{\zeta}, \bar{\pi}, \bar{v}, \bar{g}_{-l}, \bar{g}_l)\mathbf{P}_{\zeta_{n+k}, \pi_{n+k}, v_{n+k}}\left(w_{n+k} \in \frac{[\bar{a}t, \bar{\bar{a}}t]}{\lambda_{n+1} \dots \lambda_{n+k}}\right)$$

This implies that if $\mathcal{B}$ is $\mathcal{F}_{-l,l}$ measurable then

$$\mathbf{P}\left(\frac{w_n}{t} \in [\bar{a}, \bar{\bar{a}}], n \text{ is } (R,l) - \text{maximal and } \tau^n\omega \in \mathcal{B}\right) \sim \tilde{\nu}_l(\mathcal{B})\tilde{c}(\bar{a}^{-s} - \bar{\bar{a}}^{-s})t^{-s}$$

and $\tilde{\nu}_l \Rightarrow \tilde{\nu}$ as $l \to \infty$. Now the proof of (8.2) proceeds similarly to the proof of Theorem 5. To pass from (8.2) to (8.1) we note that $\frac{\mathfrak{m}_j}{w_{\bar{n}_j}}$ is well approximated by $\sum_{|k|<l} l_{n+k}(\mathbf{u}_{n+k})\Lambda_{n,k}$ provided that l is sufficiently large. Here

$$\Lambda_{n,k} = \begin{cases} \lambda_{n+1} \dots \lambda_{n+k} & \text{if } k > 0 \\ 1 & \text{if } k = 0 \\ \lambda_{n+k+1} \dots \lambda_n & \text{if } k < 0 \end{cases}.$$

Accordingly, in the limit $l \to \infty$, we have $\mathfrak{m}_j = w_{\bar{n}_j}\mathcal{H}(\omega^{(j)})$ for some measurable function $\mathcal{H}$. Now Lemma B.1 shows that (8.2) implies (8.1) with $d\tilde{\nu} = \mathcal{H}^s d\tilde{\nu}^*$.

8.2. **Arithmetic case.** We note that condition (2.6) was used in Section 7 to show that (7.2) does not have solutions. On the other hand if (7.2) has a non-trivial solution then the analysis of Appendix C has to be modified. Namely, the non-arithmetic local limit theorem (Lemma C.3) has to be replaced by its arithmetic version. This will cause replacing t^{-s} in the estimates of Theorem 9 by $t^{-s}g(\{\ln(t/\Delta)\})$ where $\{\dots\}$ denotes the fractional part, Δ is the step of the progression containing the distribution of $\ln\lambda$ and g is some continuous function. As a result the estimates of Section 5 have to be replaced by

$$\mathbf{P}(\rho_n > t) \sim t^{-s}g_1(\{\ln(t/\Delta)\}), \quad \mathbf{P}(\mathfrak{m}_j > t) \sim t^{-s}g_2(\{\ln(t/\Delta)\}).$$

Thus there would exists a measure μ on $\mathbb{R}^+$ such that $\mu([t,\infty)) = \bar{g}(\{\ln(t/\Delta)\})t^{-s}$ and the limit points of the distribution of the normalized hitting times will be of the form

$$\sum_j \Theta_j(\Gamma_j - 1)$$

where Γ_j are iid mean 1 exponential random variables and Θ_j is a Poisson process with measure $\mu_{\bar{\Delta}}$ for some $0 \le \bar{\Delta} < \Delta$ where $\mu_{\bar{\Delta}}(A) = \mu(e^{\bar{\Delta}}A)$.

In particular, we would like to note that regardless of condition (2.6) we always have

$$\mathbf{P}(\rho_n > t) \le Ct^{-s}. \tag{8.3}$$

Appendix A. Occupation times for Markov chains.

We recall two facts about general Markov chains with discrete state space. Within this section the notation $\mathbb{P}$ will be used for probabilities concerned with Markov chains and $\mathbb{E}$ for the corresponding expectations. First, the number of visits to a given state conditioned on the event that this state will be visited has a geometric distribution. Second, consider a Markov chain with transition probabilities p_{ij}. Let $\tilde{p}_{jk}$ be the probability that the chain starting at j ever visits k. Condition the chain on having at least one visit to k. Then before coming to k the chain evolves as a Markov chain with transition probabilities

$$p^*_{ij} = \frac{p_{ij}\tilde{p}_{jk}}{\sum_r p_{ir}\tilde{p}_{rk}}. \tag{A.1}$$

Similarly if we condition the chain on never visiting k then the chain evolves as a Markov chain with transition probabilities

$$p^{**}_{ij} = \frac{p_{ij}(1-\tilde{p}_{jk})}{\sum_r p_{ir}(1-\tilde{p}_{rk})}. \tag{A.2}$$

We shall now use these facts to analyze the joint distribution for the number of visits to different sites.

Namely let a and b be two states of a transient chain Z such that

$$p^{n_0}_{ab} > \varepsilon, \quad p^{n_0}_{ba} > \varepsilon, \tag{A.3}$$

where $p^{n_0}_{ab}$ denotes the transition probability after n_0 steps.

Let q_a (q_b) denote the probability that a (respectively b) is visited at least once and p_a (p_b) denote the probability that the chain started from a (respectively b) does not return to that state again. Let ξ_a (ξ_b) be the number of visits to a (respectively b).

It is useful to consider our Markov chain only at the times when it visits either a or b. Denote the resulting chain by $\tilde{Z}$. This chain is governed by the transition matrix

$$\begin{pmatrix} r_{aa} & r_{ab} & r_{ac} \\ r_{ba} & r_{bb} & r_{bc} \\ 0 & 0 & 1 \end{pmatrix}, \tag{A.4}$$

where c denotes the absorbing state which is reached by the particle after the last visit of the set $\{a, b\}$. Denote $\tilde{\varepsilon} = \varepsilon/n_0$. Since $\tilde{Z}$ is obtained from Z by skipping some states, (A.3) implies that

$$\mathbb{P}(\tilde{Z} \text{ visits } b \text{ before } n_0 | \tilde{Z}_0 = a) \geq \varepsilon.$$

(Note that here the time which is $\leq n_0$ is that of the Markov chain $\tilde{Z}$, not Z.) On the other hand this probability is equal to

$$\sum_{j=0}^{n_0-1} r^j_{aa} r_{ab} \leq n_0 r_{ab}.$$

This implies that

$$r_{ab} \geq \tilde{\varepsilon} \text{ and by symmetry } r_{ba} \geq \tilde{\varepsilon}. \tag{A.5}$$

Lemma A.1. *Suppose that* (A.3) *is satisfied. Then there is a constant C depending on $n_0, \varepsilon > 0$ such that for all initial conditions of the chain satisfying*

$$q_a > \varepsilon, \quad q_b > \varepsilon \tag{A.6}$$

the following holds:

$$\mathrm{Corr}(\xi_a, \xi_b) > 1 - \frac{C}{\mathbb{E}(\xi_a)}. \tag{A.7}$$

Remark. It is implicit in the statement of the Lemma that the constant C above is uniform over all Markov chains satisfying (A.3) and (A.6).

Remark. Remembering that

$$\mathbb{E}(\xi_a) = \frac{q_a}{p_a}, \tag{A.8}$$

we see that (A.7) is equivalent to

$$\mathrm{Corr}(\xi_a, \xi_b) > 1 - \bar{C} p_a \tag{A.9}$$

where $\bar{C}$ depends only on ε and n_0. Therefore this estimate is only interesting if p_a is very small. In this paper we apply Lemma A.1 to traps where the walker spends a lot of time so that the probability of not returning to a site in a trap is small.

Of course by symmetry we also have

$$\mathrm{Corr}(\xi_a, \xi_b) > 1 - \bar{C} p_b \tag{A.10}$$

However it is not difficult to see directly that p_a and p_b are of the same order. Namely we have

$$p_a = r_{ac} + r_{ab} \sum_{j=0}^{\infty} r_{bb}^j r_{bc} = r_{ac} + r_{ab}(1 - r_{bb})^{-1} r_{bc}$$

and similarly $p_b = r_{bc} + r_{ba}(1 - r_{aa})^{-1} r_{ac}$. It is now obvious that

$$p_a \le r_{ac} + (1 - r_{bb})^{-1} r_{bc} \le r_{ac} + \frac{r_{bc}}{\tilde{\varepsilon}} \text{ and } p_b \ge r_{bc} + r_{ba} r_{ac} \ge r_{bc} + \tilde{\varepsilon} r_{ac}.$$

This implies that

$$p_a/p_b \le \tilde{\varepsilon}^{-1} \text{ and by symmetry } p_b/p_a \le \tilde{\varepsilon}^{-1}. \tag{A.11}$$

Proof. We have $\xi_b = U + V + W$, where U is the number of visits to b before the first visit to a, W is the number of visits to b after the last visit to a and $V = \sum_{j=1}^n V_j$, where V_j is the number of visits to b between j-th and $j+1$-st visit to a. Then V_j are iid. Let $v_{ab} = \mathbb{E}(V_j)$. We claim that the following uniform bounds hold

$$v_{ab} = \mathcal{O}(1), \quad \mathbb{E}(U) = \mathcal{O}(1), \quad \mathbb{E}(W) = \mathcal{O}(1)$$

where the implied constants depend only on ε and n_0. Indeed, by (A.1)

$$\mathbb{P}(V_j = l) = r^*_{ab} (r^*_{bb})^l r^*_{ba}$$

so the estimate of v_{ab} follows from the fact that $r^*_{bb} = 1 - r^*_{ba} < 1 - r_{ba} < 1 - \tilde{\varepsilon}$. The estimate of $\mathbb{E}(U)$ is the same and the estimate of $\mathbb{E}(V)$ is similar except that we have to use (A.2) instead of (A.1).

Therefore

$$\mathbb{E}(\xi_b | \xi_a = k+1) = k v_{ab} + \mathcal{O}(1). \tag{A.12}$$

Hence

$$\mathbb{E}(\xi_b) = v_{ab} \mathbb{E}(\xi_a) + \mathcal{O}(1) \text{ and by symmetry } \mathbb{E}(\xi_a) = v_{ba} \mathbb{E}(\xi_b) + \mathcal{O}(1), \tag{A.13}$$

$$\mathbb{E}(\xi_a \xi_b) = v_{ab} \mathbb{E}(\xi_a^2) + \mathcal{O}(\mathbb{E}(\xi_a)).$$

Combining the last two equalities we obtain

$$\mathrm{Cov}(\xi_a, \xi_b) = v_{ab} \mathrm{Var}(\xi_a) + \mathcal{O}(\mathbb{E}(\xi_a))$$

$$= v_{ab} \mathrm{Var}(\xi_a) \left(1 + \mathcal{O}\left(\frac{1}{\mathbb{E}(\xi_a)}\right)\right) \tag{A.14}$$

where the last step uses (A.8) and

$$\mathrm{Var}(\xi_a) = \frac{q_a(2 - q_a - p_a)}{p_a^2}. \tag{A.15}$$

It follows from (A.8) and (A.11) that

$$n_0^{-1} \varepsilon^2 \le \mathbb{E}(\xi_a) \mathbb{E}(\xi_b)^{-1} \le n_0 \varepsilon^{-2}. \tag{A.16}$$

Therefore interchanging roles of a and b in (A.14) we get

$$\text{Cov}(\xi_a,\xi_b)=v_{ba}\text{Var}(\xi_b)\left(1+\mathcal{O}\left(\frac{1}{\mathbb{E}(\xi_a)}\right)\right). \tag{A.17}$$

Multiplying the two expressions in (A.13) we obtain

$$v_{ab}v_{ba}=1+\mathcal{O}\left(\frac{1}{\mathbb{E}(\xi_a)}\right). \tag{A.18}$$

Finally, multiplying (A.14) and (A.17) and using (A.18) we get

$$\frac{\text{Cov}^2(\xi_a,\xi_b)}{\text{Var}(\xi_a)\text{Var}(\xi_b)}=1+\mathcal{O}\left(\frac{1}{\mathbb{E}(\xi_a)}\right).$$

□

Let $\|\dots\|$ denote the L^2 norm.

Corollary A.2. *Under the conditions of Lemma A.1 there exists a constant $\bar{C}$ (depending only on ε and n_0) such that*

$$\left\|\frac{\xi_a}{\mathbb{E}(\xi_a)}-\frac{\xi_b}{\mathbb{E}(\xi_b)}\right\|\le\bar{C}\left[\frac{1}{\sqrt{\mathbb{E}(\xi_a)}}+(1-q_a)+(1-q_b)\right].$$

Proof. We have

$$\left\|\frac{\xi_a}{\mathbb{E}(\xi_a)}-\frac{\xi_b}{\mathbb{E}(\xi_b)}\right\|=\left\|\frac{\xi_a-\mathbb{E}(\xi_a)}{\mathbb{E}(\xi_a)}-\frac{\xi_b-\mathbb{E}(\xi_b)}{\mathbb{E}(\xi_b)}\right\|$$
$$=\left\|\frac{\xi_a-\mathbb{E}(\xi_a)}{\sqrt{\text{Var}(\xi_a)}}\frac{\sqrt{\text{Var}(\xi_a)}}{\mathbb{E}(\xi_a)}-\frac{\xi_b-\mathbb{E}(\xi_b)}{\sqrt{\text{Var}(\xi_b)}}\frac{\sqrt{\text{Var}(\xi_b)}}{\mathbb{E}(\xi_b)}\right\|$$
$$\le\left\|\frac{\xi_a-\mathbb{E}(\xi_a)}{\sqrt{\text{Var}(\xi_a)}}-\frac{\xi_b-\mathbb{E}(\xi_b)}{\sqrt{\text{Var}(\xi_b)}}\right\|+\left|\frac{\sqrt{\text{Var}(\xi_a)}}{\mathbb{E}(\xi_a)}-1\right|+\left|\frac{\sqrt{\text{Var}(\xi_b)}}{\mathbb{E}(\xi_b)}-1\right|$$

Notice that the first term equals to $\sqrt{2(1-\text{Corr}(\xi_a,\xi_b))}$ and so it can be estimated by Lemma A.1 while the last two terms can be estimated by (A.8) and (A.15). □

To use the above corollary we need to estimate $1-q_a$. To this end we observe that if (A.3) holds then

$$\mathbb{P}_b(\xi_a=0)\le\sum_{l=0}^{\infty}r_{bb}^l r_{bc}=\frac{r_{bc}}{1-r_{bb}}\le\frac{p_b}{r_{ba}}\le\frac{p_b}{\tilde{\varepsilon}}. \tag{A.19}$$

Appendix B. Poisson process and stable distributions.

Let (X,μ) be a measure space. Recall that a Poisson process is a point process with values in X such that if $N(A)$ is the number of points in $A\subset X$ then $N(A_1),N(A_2)\dots N(A_k)$ are mutually independent if $A_1,A_2\dots A_k$ are disjoint and $N(A)$ has the Poisson distribution with parameter $\mu(A)$. If $X\subset\mathbb{R}^d$ and μ has density f with respect to the Lebesgue measure we say that f is the intensity of the Poisson process.

Lemma B.1. *(see* [19]*, sections 2.3 and 5.2)*

(a) If $\{\Theta_j\}$ is a Poisson process on X and $\psi:X\to\tilde{X}$ is a measurable map then $\tilde{\Theta}_j=\psi(\Theta_j)$ is a Poisson process. If $X=\tilde{X}=\mathbb{R}$ and ψ is invertible differentiable map then the intensity of $\tilde{\Theta}$ is

$$\tilde{f}(\theta)=f(\psi^{-1}(\theta))\left|\frac{d\psi}{d\theta}\right|^{-1}.$$

(b) Let (Θ_j,Γ_j) be a point process on $X\times Z$ such that $\{\Theta_j\}$ is a Poisson process on X and $\{\Gamma_j\}$ are Z-valued random variables which are i.i.d. and independent of $\{\Theta_k\}$ then (Θ_j,Γ_j) is a Poisson process on $X\times Z$.

(c) If in (b) $X = Z = \mathbb{R}$ then $\tilde{\Theta} = \{\Gamma_j \Theta_j\}$ is a Poisson process. Its intensity is

$$\tilde{f}(\theta) = \mathbf{E}\left(f\left(\frac{\theta}{\Gamma}\right)\frac{1}{\Gamma}\right).$$

Lemma B.2. *(see [28], Theorem 1.4.2)*

(a) If $0 < s < 1$ and Θ_j is a Poisson process with intensity $\theta^{-(1+s)}$ then $\sum_j \Theta_j$ has stable distribution of index s.

(b) If $1 < s < 2$ and Θ_j is a Poisson process with intensity $\theta^{-(1+s)}$ then

$$\lim_{\delta \to 0}\left[\left(\sum_{\delta < \Theta_j} \Theta_j\right) - \frac{1}{(s-1)\delta^{s-1}}\right]$$

has stable distribution of index s.

(c) If $s = 1$ and Θ_j is a Poisson process with intensity θ^{-2} then

$$\lim_{\delta \to 0}\left[\left(\sum_{\delta < \Theta_j} \Theta_j\right) - |\ln \delta|\right]$$

has stable distribution of index 1.

Appendix C. Renewal theorem for a system of contractions.

C.1. **Main result.** Let M_1 amd M_2 be compact metric spaces, and let $\omega = \{g_j\}$ be a sequence of iid M_2 valued random variables. Suppose that there is a C^η map $G : M_1 \times M_2 \to M_1$. Here and below C^η denotes the space of Holder maps, that is

$$d(G(\Phi_1, g_1), G(\Phi_2, g_2)) < C\left[d^\eta(\Phi_1, \Phi_2) + d^\eta(g_1, g_2)\right].$$

Note that later we impose a stronger condition on Φ dependence (see (C.1) below). Consider Markov process on M_1

$$\Phi_{j+1} = G_j(\Phi_j), \text{ where } G_j(\Phi) = G(\Phi, g_j).$$

We suppose that the maps G_j are contractions in the sense that there exist constants C and $\theta < 1$ such that if Φ_j' and Φ_j'' are two realizations of this Markov chain starting from Φ_0' and Φ_0'' and evolving in the same environment ω then with probability 1 we have

$$d(\Phi_j', \Phi_j'') < C\theta^j d(\Phi_0', \Phi_0''). \tag{C.1}$$

Denote $\tilde{\omega} = \{(g_j, \Phi_j)\}$, $\Omega = (M_2)^{\mathbb{Z}}$, $\tilde{\Omega} = (M_1 \times M_2)^{\mathbb{Z}}$. Let $\mathbb{P}$ be the distribution of ω. In view of (C.1) there is a unique stationary distribution $\tilde{\mathbb{P}}$ on $\tilde{\Omega}$ whose projection onto Ω is $\mathbb{P}$. Namely, conditioned on ω the distribution of $\{\Phi_j\}$ is delta measure concentrated at $\{\bar{\Phi}_j\}$ where $\bar{\Phi}_j$ is constructed as follows. Take $\Phi^* \in M_1$ and let $\Phi_{k,j} = G_{j-1} \dots G_{k+1} G_k(\Phi^*)$. Then $\bar{\Phi}_j = \lim_{k \to -\infty} \Phi_{k,j}$. Let $b(\Phi, g)$ be a positive C^η function on $M_1 \times M_2$ and $\Delta(\tilde{\omega})$ be a positive continuous function on $\tilde{\Omega}$. Denote $b_j = b(\Phi_j, g_j)$, $\Delta_j = \Delta(\tau^j \omega)$, where τ denotes the shift. For $k > 0$ denote

$$z_{n,k}^+ = \sum_{j=1}^{k} b_n b_{n+1} \dots b_{n+j} \Delta_{n+j}, \quad z_{n,k}^- = \sum_{j=1}^{k} b_n b_{n-1} \dots b_{n-j} \Delta_j.$$

Let

$$z_n^+ = \lim_{k \to \infty} z_{n,k}^+, \quad z_n^- = \lim_{k \to \infty} z_{n,k}.$$

Denote $a = \ln b$. Suppose that $\tilde{\mathbb{E}}(a) < 0$ but for any t there exists N such that

$$\tilde{\mathbb{P}}\left(\prod_{j=1}^{N} b_j > t\right) > 0. \tag{C.2}$$

Denote by $\mathbb{S}^1$ a set of complex numbers of absolute value 1.

Theorem 9. *Suppose that for any numbers $u, \bar{u} \in \mathbb{R}$, there exists no continuous function $\mathfrak{h} : M_1 \to \mathbb{S}^1$ such that the following equation is satisfied $\tilde{\mathbb{P}}$ almost surely*

$$e^{iua(\Phi,g)} = e^{i\bar{u}} \frac{\mathfrak{h}(\Phi)}{\mathfrak{h}(G(\Phi,g))}. \tag{C.3}$$

Then there are constants $s > 0$, $\bar{s} > s$ and $\bar{C} > 0$ such that such that

(a) If $k > \bar{C} \ln t$ then

$$\mathbb{P}(z_n^+ - z_{n,k}^+ > 1) < \bar{C} t^{-\bar{s}}, \quad \mathbb{P}(z_n^- - z_{n,k}^- > 1) < \bar{C} t^{-\bar{s}}.$$

(b) There exists a function $f(\Phi)$ such that if $k > \bar{C} \ln t$ then

$$\mathbb{P}(z_{n,k}^+ > t | \Phi_0 = \Phi) \sim f(\Phi) t^{-s}.$$

(Here and below '$\sim$' means that the estimates are uniform in k, that is given ε there exists t_0 such that if $t > t_0$ and $k > \bar{C} \ln t$ then

$$\left| t^s \mathbb{P}(z_{n,k}^+ > t | \Phi_0 = \Phi) - f(\Phi) \right| < \varepsilon.)$$

In particular

$$\mathbb{P}(z_n^+ > t | \Phi_0 = \Phi) \sim f(\Phi) t^{-s}.$$

(c) There exists a measure $\tilde{\nu}$ on $\tilde{\Omega}$ such that for any continuous function H on $\tilde{\Omega}$ the following asymptotic relations hold if $k > \bar{C} \ln t$ and $t \to \infty$:

$$\tilde{\mathbb{E}}(1_{z_{n,k}^- > t} H(\tilde{\omega})) \sim t^{-s} \tilde{\nu}(H).$$

In particular

$$\tilde{\mathbb{E}}(1_{z_n^- > t} H(\tilde{\omega})) \sim t^{-s} \tilde{\nu}(H).$$

C.2. **Renewal theorem and large deviations.** We will deduce Theorem 9 from a large deviation bound. Let $y_n = \sum_{j=0}^{n-1} a_j$.

Theorem 10. *Suppose that* (C.3) *has no solutions. Then there is a number $\alpha^* > 0$ and a strictly convex analytic function $\gamma : [\tilde{\mathbb{E}}(a), \alpha^*) \to \mathbb{R}$ such that*

$$\gamma'(\alpha) > 0 \text{ if } \alpha > \tilde{\mathbb{E}}(a), \quad \lim_{\alpha \to \alpha^*} \gamma'(\alpha) = +\infty \tag{C.4}$$

and

(a) If $\alpha > \alpha^*$ then for each $\beta \in \mathbb{R}$, $\mathbb{P}_\Phi(y_n \geq \alpha n) = \mathcal{O}(e^{-\beta n})$;

(b) If $\alpha \in (\tilde{\mathbb{E}}(a), \alpha^*)$ then there is a measure ν_α on $\tilde{\Omega}$ and a function $h_\alpha : M_1 \mapsto \mathbb{R}$ such that for any $J \subset [0, \infty]$ and any continuous function $H : \tilde{\Omega} \mapsto \mathbb{R}$

$$\mathbb{E}_\Phi \left(1_{y_n - \alpha n \in J} H(\tau^n \tilde{\omega}) \right) \sim \frac{e^{-\gamma(\alpha) n}}{\sqrt{n}} \nu_\alpha(H) h_\alpha(\Phi) \int_J e^{-\gamma'(\alpha) t} \gamma'(\alpha) dt,$$

(c) If $\alpha \in (\tilde{\mathbb{E}}(a), \alpha^*)$ then for any $J \subset [0, \infty]$ and any continuous functions $H, \hat{H} : \tilde{\Omega} \mapsto \mathbb{R}$

$$\tilde{\mathbb{E}} \left(1_{y_n - \alpha n \in J} \hat{H}(\tilde{\omega}) H(\tau^n \tilde{\omega}) \right) \sim \frac{e^{-\gamma(\alpha) n}}{\sqrt{n}} \hat{\nu}_\alpha(\hat{H}) \nu_\alpha(H) \int_J e^{-\gamma'(\alpha) t} \gamma'(\alpha) dt.$$

Theorem 10 is proven in subsection C.3. Here we use this theorem to obtain Theorem 9.

Let $\alpha_0 = \arg\min \frac{\gamma(\alpha)}{\alpha}$ and $s = \frac{\gamma(\alpha_0)}{\alpha_0}$. Note that by (C.4) the minimum is achieved strictly inside $(0, \alpha^*)$. Indeed from part (b) of Theorem 10 it is clear that γ is non-negative on $(\tilde{\mathbb{E}}(a), \alpha^*)$ and since it is strictly increasing it follows that it is in fact strictly positive on this interval. Accordingly $\lim_{\alpha \to 0+} \frac{\gamma(\alpha)}{\alpha} = +\infty$ and so there exists $\varepsilon_1 > 0$ such that the minimum can not be achived on $[0, \varepsilon_1]$. Next we show that the minimum can not be achived near α^*. Notice that γ is increasing. Now if $\lim_{\alpha\to\alpha^*}\gamma(\alpha)=+\infty$ then the statement is obvious. Otherwise the formula

$$\left(\frac{\gamma(\alpha)}{\alpha}\right)'=\frac{\gamma'(\alpha)}{\alpha}-\frac{\gamma(\alpha)}{\alpha^2}$$

shows that there exists some $\varepsilon_2>0$ such that $\frac{\gamma(\alpha)}{\alpha}$ is increasing on $(\alpha^*-\varepsilon_2,\alpha^*)$ proving our claim that $s\in(0,\alpha^*)$.

Lemma C.1. *There exist constants C and $\tilde{s}>0$ such that*

(a) for each Φ we have $\mathbb{P}_\Phi(\max y_n\geq Y)\leq Ce^{-sY}$;

(b) For each n we have $\mathbb{P}_\Phi(\exists l>k: y_n\geq Y$ and $y_{n+l}\geq Y)\leq\frac{C}{\sqrt{Y}}e^{-(sY+\tilde{s}k)}$. In particular, $\mathbb{P}_\Phi(\exists n_1,n_2: n_2>n_1+k$ and $y_{n_j}\geq Y$ for $j=1,2)\leq Ce^{-(sY+\tilde{s}k)}$.

Proof. By Theorem 10

$$\mathbb{P}_\Phi(y_n\geq Y)\leq\frac{C}{\sqrt{n}}\exp\left[-\left(\frac{\gamma(Y/n)}{Y/n}\right)Y\right]$$

$$\leq\frac{C}{\sqrt{n}}\exp[-sY]\exp\left[-c\left(\frac{Y}{n}-\alpha_0\right)^2Y\right]. \tag{C.5}$$

The main contribution to this sum comes from $n\approx Y/\alpha_0$. For those n

$$c\left(\frac{Y}{n}-\alpha_0\right)^2Y\leq\tilde{c}\frac{(Y-\alpha_0 n)^2}{Y}.$$

Since

$$\sum_n\frac{1}{\sqrt{Y}}\exp\left[-\tilde{c}\frac{(Y-\alpha_0 n)^2}{Y}\right]\leq C$$

part (a) is proved.

To prove (b) observe that by Markov property

$$\mathbb{P}_\Phi(y_n\geq Y\text{ and }y_{n+l}\geq Y)\leq\frac{C}{\sqrt{Y}}e^{-(sY+\tilde{s}k)}\leq\mathbb{P}_\Phi(y_n\geq Y)\max_{\bar\Phi}\mathbb{P}_{\bar\Phi}(y_l>0).$$

The second term is less than $\frac{C}{\sqrt{l}}e^{-\gamma(0)l}$ due to Theorem 10 while the first term is less than $\frac{C}{\sqrt{Y}}e^{-sY}$ by part (a). Now the first inequality of part (b) follows by summation over $l>k$ and the second one follows by summation over n and l. □

Lemma C.2. *Suppose that $n,Y\to\infty$ so that $\frac{n-\frac{Y}{\alpha_0}}{\sqrt{Y}}\to\beta$ Denote*

$$\Omega_n=\{y_n\geq Y, y_m<Y\textit{ for all }0<m<n\}.$$

Then for each Φ and a continuous function $\hat{H}:\mathbb{R}\times\tilde\Omega\to\mathbb{R}$ the following limits exist

(a) $\lim_{n\to\infty}\mathbb{P}_\Phi(\Omega_n)\sqrt{Y}e^{sY}$;

(b) $\lim_{n\to\infty}\mathbb{E}_\Phi(1_{\Omega_n}\hat{H}(y_n-Y,\tau^n\tilde\omega)\sqrt{Y}e^{sY}$.

Moreover both limits are bounded by $\mathrm{Const}e^{-c\beta^2}$.

Proof. (a) By Lemma C.1(b) it is enough to show that for each k the following limit exists

$$\lim_{n\to\infty}\mathbb{P}_\Phi(y_n\geq Y, y_m<Y\text{ for }n-k<m<n)\sqrt{Y}e^{sY}. \tag{C.6}$$

The limiting expression equals to

$$\int\mathbb{P}_\Phi(y_{n-k}\in[Y-z_1,Y-z_2])dP_k^\Phi(z_1,z_2)\sqrt{Y}e^{sY}$$

where P_k^Φ is the distribution function of the random vector $(y_k, \max_{0<j\leq k} y_j)$ for our Markov chain started from Φ. The integral above is the limit of Lebesgue-Stiltjes sums where each term has form

$$\mathbb{P}_\Phi(y_{n-k} \in [Y - z_1, Y - z_2])\mathbb{P}_{\Phi_{n-k}}(y_k \in [z_1, z_1 + \varepsilon), \max_{0<j\leq k} y_j \in [z_2, z_2 + \varepsilon))\sqrt{Y} e^{sY}.$$

Since for each z_1, z_2 the last probability is a function of Φ_{n-k} part (a) follows from Theorem 10.

For part (b) it is sufficient to restrict our attention to a dense set of functions $\hat{H}$. In particular we can assume that $\hat{H}$ depends only on finitely many coordinates, that is we need to compute the limit

$$\lim_{n\to\infty} \mathbb{E}_\Phi(1_{y_n\geq Y, y_m<Y \text{ for } n-k<m<n} \tilde{H}(\Phi_{n-k}, g_{n-k} \dots \Phi_{n+k}, g_{n+k}))\sqrt{Y} e^{sY}.$$

for some $\tilde{H} : (M_1 \times M_2)^{2k+1} \to \mathbb{R}$. The analysis of the last limit is the same as the analysis of (C.6).

Finally the fact that above limits are $\mathcal{O}\left(e^{-c\beta^2}\right)$ follows from the estimates in the proof of Lemma C.1 (see (C.5)). □

Proof of Theorem 9. Let $Y = \ln t$. Note that $z_n^+ - z_{n,k}^+ = \sum_{j=k+1}^\infty \Delta_j e^{y_j}$. Take small $\varepsilon > 0$. Then if t is sufficiently large, the inequality $z_n^+ - z_{n,k}^+ \geq 1$ implies that there is $j > k$ such that $y_j > (1-\varepsilon)Y - \varepsilon j$. Note that $-\varepsilon j < (1-\varepsilon)Y - \varepsilon j < \varepsilon j$ provided $\bar{C}$ is large enough. Hence

$$\mathbb{P}_\Phi(y_j > (1-\varepsilon)Y - \varepsilon j) \leq \frac{C}{\sqrt{j}} e^{-\bar{\gamma} j}$$

where $\bar{\gamma} = \min_{[-\varepsilon,\varepsilon]} \gamma$. Summing over j we get

$$\mathbb{P}_\Phi(z_n^+ - z_{n,k}^+ > 1) \leq C(\varepsilon) e^{-\bar{\gamma} k}.$$

Since $k \geq \bar{\Delta} \ln t$ and $\bar{\Delta}$ can be taken sufficiently large this proves the first inequlity of part (a). The proof of the second inequlity is similar.

To prove part (b) take $M \gg 1$. We claim that terms with $y_{-n} \leq Y - M$ can be ignored. Indeed for terms with $Y - M - 1 < y_n < Y - M$ to make a contribution greater than $e^{-M/2}$ there should be at least $e^{M/2}/C$ such terms. By Lemma C.1 the probability of such an event is

$$\mathcal{O}\left(\exp -[s(Y-M) + \tilde{s}\exp(M/2)]\right)$$

which establishes our claim. Therefore for large M and l we can approximate $\mathbb{P}_\Phi(z_{n,k}^+ \geq t)t^s$ and $\mathbb{P}_\Phi(z_n^+ \geq t)t^s$ by $\sum_n \mathbb{P}_\Phi(1_{\Omega_{n,M}} 1_{B_{n,l,\varepsilon}})t^s$ where

$$\Omega_{n,M} = \{y_{-n} \geq Y - M, y_{-m} < Y - M \text{ for } 0 \leq m < n\},$$

$$B_{n,l} = \left\{ \sum_{|m-n|<l} \Delta_m e^{y_m} \geq t - \varepsilon \right\}.$$

The fact that the last $\sum_n \mathbb{P}_\Phi(1_{\Omega_{n,M}} 1_{B_{n,l,\varepsilon}})t^s$ approaches the limit as $t \to \infty$ follows from Lemma C.2(b). This proves part (b). The proof of part (c) is similar. □

C.3. **Large deviations.** We follow the approach of [32].

To simplify the notation we assume for the rest of this section that (C.1) holds with $C = 1$, the general case can be reduced to this one by considering our Markov chains only at the times which are multiples of a sufficiently large n_0.

For $\kappa \geq 0$ consider operators P_κ given by

$$P_\kappa(h)(\Phi) = \mathbb{E}_\Phi(e^{\kappa a(\Phi, g)} h(\Phi_1)).$$

P_κ is a positive operator preserving the space of C^η functions. Moreover it has many invariant cones as we describe below. Let

$$\mathcal{C}_K = \{h \geq 0 : \text{ for all } \tilde{\Phi}, \tilde{\tilde{\Phi}} \text{ we have } h(\tilde{\Phi}) \leq e^{K d^\eta(\tilde{\Phi}, \tilde{\tilde{\Phi}})} h(\tilde{\tilde{\Phi}})\}.$$

A direct computation (using (C.1) with $C=1$) shows that $P_\kappa(\mathcal{C}_K)\subset\mathcal{C}_{\bar K}$ where $\bar K=K\theta^\eta+\kappa\mathbf{H}$ and $\mathbf{H}$ is the Holder constant of a with respect to Φ variable. Now [22] shows that if K is so large that $K>\bar K$ then P_κ contracts the Hilbert metric on $\mathcal{C}_K$ and so there exist positive eigenfunctions

$$P_\kappa h_\kappa = e^{\lambda_\kappa} h_\kappa \tag{C.7}$$

in $\mathcal{C}_K$ and moreover for any two elements h', h'' of $\mathcal{C}_K$ the directions of $P_\kappa^n h'$ and $P_\kappa^n h''$ converge to each other exponentially fast. This in turn implies that the rest of the spectrum of P_κ is contained in a disc of radius strictly smaller than e^{λ_κ}. Since e^{λ_κ} is an isolated eigenvalue of P_κ, λ_κ depends analytically on κ.

We need the fact that the map $\kappa\to\lambda_\kappa$ is strictly convex. To see this we need formulas for derivatives of λ with respect to κ. To this end let ν_κ be the eigenvector of the adjoint operator

$$\nu_\kappa(P_\kappa h)=e^{\lambda_\kappa}\nu_\kappa(h).$$

Differentiating (C.7) we get

$$\mathbb{E}_\Phi\left(a(\Phi,g)e^{\kappa a(\Phi,g)}h_\kappa(G(\Phi,g))\right)+P_\kappa(h_\kappa')=\lambda' e^{\lambda_\kappa}h_\kappa+e^{\lambda_\kappa}h_\kappa'.$$

Applying ν_κ to both sides we get

$$\lambda_\kappa=\frac{\nu_\kappa(\mathbb{E}(a(\Phi,g)e^{\kappa a(\Phi,g)}h_\kappa(G(\Phi,g))))}{\nu_\kappa(e^{\lambda_\kappa}h_\kappa)}. \tag{C.8}$$

Let

$$\tilde a_\kappa=\kappa a-\lambda_\kappa+\ln h_\kappa-\ln h_\kappa(G(\Phi,g)).$$

Then

$$\mathbb{E}_\Phi\left(e^{\tilde a_\kappa}\right)=1 \tag{C.9}$$

so we can consider a Markov chain with generator

$$\tilde P_\kappa(h)(\Phi)=\mathbb{E}_\Phi\left(e^{\tilde a_\kappa}h(\Phi_1)\right).$$

Observe that

$$\tilde P_\kappa=e^{-\lambda_\kappa}M_\kappa^{-1}P_\kappa M_\kappa$$

where M_κ denotes the multiplication by $e^{\tilde a_\kappa}$ so the eigenvector of the adjoint operator (which is the stationary measure for our Markov process) equals

$$m_\kappa(h)=\nu_\kappa(hh_\kappa).$$

Normalize m_κ by the condition $m_\kappa(1)=1$. Then m_κ is the invariant measure for the Markov process with transition operator P_κ. Denoting by $\mathbf{m}_\kappa$ the corresponding invariant measure on $\tilde\Omega$ we can rewrite (C.8) as

$$\lambda_\kappa'=\mathbf{m}_\kappa(a). \tag{C.10}$$

Next we compute λ_κ''. Fix a κ_0 and let

$$\bar P_\kappa h=\mathbb{E}_\Phi(e^{\bar a_\kappa}h(\Phi_1))$$

where

$$\bar a_\kappa=\kappa(a-\mathbf{m}_{\kappa_0}(a))+\ln h_{\kappa_0}-\ln h_{\kappa_0}(G(\Phi,g))-\lambda_{\kappa_0}+\kappa_0\mathbf{m}_{\kappa_0}(a).$$

Then the leading eigenvalue of $\bar P_\kappa$ is

$$\bar\lambda_\kappa=\lambda_\kappa-(\kappa-\kappa_0)\mathbf{m}_{\kappa_0}(a)-\lambda_{\kappa_0}$$

and so $\bar\lambda''(\kappa_0)=\lambda''(\kappa_0)$. Let $\bar h_\kappa$ be the leading eigenvector of $\bar P_\kappa$ and $\bar\mu_k$ be the leading eigenvalue of the adjoint operator. Then we have

$$\bar P_\kappa'' h+2\bar P_\kappa'\bar h_\kappa'+P_\kappa\bar h_\kappa''=e^{\bar\lambda_\kappa}(\bar\lambda_\kappa')'\bar h_\kappa+e^{\bar\lambda_\kappa}\bar\lambda_\kappa''\bar h_\kappa+2e^{\bar\lambda_\kappa}\bar\lambda_\kappa'\bar h_\kappa'+e^{\bar\lambda_\kappa}\bar h_\kappa''.$$

Applying $\bar\nu_\kappa$ to both sides and using that

$$\bar\lambda_{\kappa_0}=0,\quad \bar\lambda_{\kappa_0}'=0,\quad \bar h_{\kappa_0}=1,\quad \bar\nu_{\kappa_0}=m_{\kappa_0}$$

we get

$$\lambda''_{\kappa_0} = \mathbf{m}_{\kappa_0}(\hat{a}^2) + 2\mathbf{m}_{\kappa_0}(\hat{a}(\Phi_0, g_0)\bar{h}_{\kappa_0}(\Phi_1))$$

where $\hat{a} = a - \mathbf{m}_{\kappa_0}(a)$. Applying the same argument to P_κ^n, which has the leading eigenvalue $e^{n\lambda_\kappa}$ we get

$$n\lambda''_{\kappa_0} = \mathbf{m}_{\kappa_0}\left(\left(\sum_{j=0}^{n-1} \hat{a}(\Phi_j, g_j)\right)^2\right) + 2\mathbf{m}_{\kappa_0}\left(\left(\sum_{j=0}^{n-1} \hat{a}(\Phi_j, g_j)\right)\bar{h}_{\kappa_0}(\Phi_n)\right).$$

Since the Markov process with transition operator $\tilde{P}_\kappa$ has a spectral gap the measure $\mathbf{m}_\kappa$ is ergodic and hence

$$\frac{1}{n}\sum_{j=0}^{n-1} \hat{a}(\Phi_j, g_j) \to 0$$

almost surely. Therefore

$$\lambda''_{\kappa_0} = \lim_{n\to\infty} \frac{1}{n}\mathbf{m}_{\kappa_0}\left(\left(\sum_{j=0}^{n-1} \hat{a}(\Phi_j, g_j)\right)^2\right). \tag{C.11}$$

Since the RHS of the last expression is positive we conclude that λ_κ is convex. We now show following the argument of Theorem 12 of [2] that λ_κ is actually strictly convex. Consider the following operator on $C^\eta(M_1 \times M_2)$

$$(\hat{P}_\kappa h)(\Phi_0, g_0) = \tilde{\mathbb{E}}^\kappa_{\Phi_0, g_0} h(\Phi_1, g_1).$$

Denote $\Gamma = (1-\hat{P}_\kappa)^{-1}\hat{a} = \sum_{j=0}^\infty \hat{P}_\kappa^j \hat{a}$. Then a direct computation shows that the RHS of (C.11) equals to $\mathbf{m}_\kappa(\Gamma^2 - (\hat{P}_\kappa\Gamma)^2)$. Hence if $\lambda''_{\kappa_0} = 0$ then we have

$$\mathbf{m}_\kappa((\hat{P}_{\kappa_0}\Gamma)^2) = \mathbf{m}_{\kappa_0}(\Gamma^2).$$

Since $\mathbf{m}_\kappa$ is stationary for $\hat{P}_\kappa$ this implies that

$$\mathbf{m}_{\kappa_0}((\hat{P}_{\kappa_0}\Gamma)^2) = \mathbf{m}_{\kappa_0}(\hat{P}_{\kappa_0}(\Gamma^2)).$$

Now Jensen inequality tells us that $\Gamma(\Phi, g)$ is actually independent of g, $\Gamma = \Gamma(\Phi)$. Then

$$\hat{a}(\Phi, g) = \Gamma(\Phi) - (\hat{P}_{\kappa_0}\Gamma)(\Phi, g) = \Gamma(\Phi) - \Gamma(\Phi g)$$

contradicting (C.2) (as well as (C.3)). This proves that λ_κ is strictly convex.

Let $\alpha^* = \lim_{\kappa\to+\infty} \lambda'_\kappa$. This limit exists since λ'_κ is increasing and is finite since $\lambda_\kappa \le \kappa||a||_{C^0}$. We now prove Theorem 10 with this value of α^*.

Proof. To prove part (a) we iterate (C.9) to get

$$\mathbb{E}_\Phi\left(e^{\tilde{y}_n}\right) = 1$$

where

$$\tilde{y}_n = \xi y_n - n\lambda_\kappa + \ln h(\Phi_0) - \ln h(\Phi_n).$$

Hence by Markov inequality

$$\mathbb{P}_\Phi(y_n > n\alpha) \le Ce^{-n(\kappa\alpha - \lambda_\kappa)} \le Ce^{n\kappa(\alpha^* - \alpha)}$$

where the last inequality uses that $\lambda_\kappa < \kappa\alpha^*$. This proves part (a).

To prove part (b) suppose that κ is such that $m_\kappa(a) = \alpha$. Note that if $m_0(a) = \tilde{\mathbb{E}}(a)$ so if $\alpha > \tilde{\mathbb{E}}(a)$ then κ is strictly positive.

Let $\tilde{\mathbb{E}}^\kappa_\Phi$ denote the expectation with respect to the Markov process with generator $\tilde{P}_\kappa$.

Lemma C.3. *If* (C.3) *has no solutions then there exists a function $\phi(\Phi)$ such that*

$$\tilde{\mathbb{E}}^\kappa_\Phi\left(1_{y_n - n\alpha\in I} H(\tau^n\tilde{\omega})\right)\sqrt{n} \to \mathrm{Leb}(I)\phi(\Phi)\nu(H). \tag{C.12}$$

The proof of this Lemma is given in subsection C.4.

Now take $I=[t,t+\varepsilon]$ then the RHS equals $\varepsilon\phi(\Phi)\nu(H)$ while the LHS equals

$$\sqrt{n}h_\kappa(\Phi)\mathbb{E}_\Phi\left(e^{n(\alpha\kappa-\lambda_\kappa)}\frac{H(\tau^n\tilde{\omega})}{h_\kappa(\Phi_n)}e^{\kappa t}\right)(1+o_{\varepsilon\to 0}(1)).$$

Dividing J into the segments of length $\varepsilon \ll 1$ we obtain part (b). Part (c) follows from part (b) and the Markov property. Finally observe that

$$\gamma(\alpha)=\alpha\kappa-\lambda_\kappa$$

where κ satisfies $\lambda'_\kappa=\alpha$. Thus

$$\frac{\partial\gamma}{\partial\alpha}=(\alpha-\lambda'_\kappa)\frac{\partial\kappa}{\partial\alpha}+\kappa=\kappa.$$

This proves (C.4). □

C.4. **Local Limit Theorem.** Consider

$$\tilde{P}_{\kappa,u}(h)(\Phi)=\mathbb{E}_\Phi\left(e^{\tilde{a}_\kappa-iu(a_\kappa-\alpha)}h(\Phi_1)\right).$$

Then

$$\tilde{P}^n_{\kappa,u}(h)(\Phi)=\mathbb{E}_\Phi\left(\exp\left[\sum_{j=0}^{n-1}\tilde{a}_\kappa-iu(a_\kappa-\alpha)(\Phi_j,g_j)\right]h(\Phi_n)\right).$$

Since

$$\left|\sum_{j=0}^{n-1}\left[(\tilde{a}_\kappa-iu(a_\kappa-\alpha))(\tilde{\Phi}_j,g_j)-(\tilde{a}_\kappa-iu(a_\kappa-\alpha))(\tilde{\tilde{\Phi}}_j,g_j)\right]\right|$$
$$\leq C\sum_{j=0}^{n-1}d^\eta(\tilde{\Phi}_j,\tilde{\tilde{\Phi}}_j)\leq \tilde{C}d^\eta(\tilde{\Phi}_0,\tilde{\tilde{\Phi}}_0)$$

we get

$$\|P^n_{\kappa,u}h\|_{C^\eta}\leq\left[\|h\|_{C^0}+C(u)\theta^n\|h\|_{C^\eta}\right]. \tag{C.13}$$

Using the theory of Doeblin-Fortet operators ([22]) we conclude that for each $\theta'>\theta$ the spectrum of $P_{\kappa,u}$ outside the disc of radius θ' consists of a finite number of eigenvalues with absolute values at most 1. We claim that in fact there are no eigenvalues of absolute value 1. Indeed let $e^{i\bar{u}}$ be such an eigenvalue and $\mathfrak{h}$ be the corresponding eigenfunction. Then

$$\mathbb{E}_\Phi(e^{\tilde{a}_\kappa+iua(\Phi)}\mathfrak{h}(\Phi_1))=e^{i\bar{u}}\mathfrak{h}(\Phi). \tag{C.14}$$

Let $\Phi^*=\arg\max|\mathfrak{h}|$. Without loss of generality we can assume that $|\mathfrak{h}(\Phi^*)|=1$. Now (C.9) implies that (C.14) is only possible if $|\mathfrak{h}(G(\Phi^*))|=1$ with probability 1. Iterating we see that for all n

$$|\mathfrak{h}(G_n\dots G_1(\Phi^*))|=1. \tag{C.15}$$

We claim that this implies that

$$|\mathfrak{h}(\Phi)|\equiv 1 \tag{C.16}$$

on the support of $\tilde{\mu}$. Indeed if $|\mathfrak{h}|<1-\varepsilon$ on a relatively open subset U of $\mathrm{supp}(\tilde{\mu})$ then there would exist $\bar{\Phi}\in M_1$ and $n_k\to\infty$ and such that $G_{n_k}\dots G_1(\bar{\Phi})\in U$ with positive probability. Since $G_n\dots G_1$ contracts with speed θ^n for large k we would have $|\mathfrak{h}(G_{n_k}\dots G_1(\Phi^*))|<1-\varepsilon/2$ with positive probability, contradiciting (C.15). Now (C.16) and (C.9) show that

$$e^{iua(\Phi,g)}\mathfrak{h}(G(\Phi,g))=e^{i\bar{u}}\mathfrak{h}(\Phi)$$

which contradicts (C.3).

We are now ready to prove Lemma C.3. Since the LHS of (C.12) is monotone function of H it suffices to prove the result for a dense set of functions. In particular we may assume that H depends only on finitely many coordinates

$$H=\tilde{H}(\omega_{-k+1},\dots,\omega_0,\dots\omega_{k-1}).$$

Then

$$\sqrt{n}\tilde{\mathbb{E}}_\Phi^\kappa\left(1_{y_n - n\alpha\in I} H(\tau^n\tilde\omega)\right) = \sqrt{n}\int \tilde{\mathbb{E}}_\Phi^\kappa\left(1_{y_{n-k}-(n-k)\alpha - Z - k\alpha\in I} h(\Phi_{n-k}, Z)\right) dP_k(Z)$$

where $h(\Phi, Z) = \mathbb{E}_{\Phi_0=\Phi}(\tilde H(\tau^{-k}\omega)|y_k = Z)$ and $P_k(z)$ is the distribution function of y_k. Observe that for each Z the RHS has the same form as the LHS of (C.12) except that n is replaced to $n-k$ and h depends only on one coordinate. Therefore it suffices to prove (C.12) in the case where $H = h(\Phi_0)$.

Let $\Gamma_\theta(y) = \frac{1}{\pi}\frac{1-\cos(\delta y)}{\delta y^2} e^{i\theta y}$. Then $\hat\Gamma_0(u) = (1-\frac{|u|}{\delta})_+$ and $\hat\Gamma_\theta(u) = \hat\Gamma_0(u+\theta)$. By Section 2.5 of [7] it suffices to show that for each θ, δ we have

$$\sqrt{n}\tilde{\mathbb{E}}_\Phi^\kappa(\Gamma_\theta(y_n - \alpha n)h(\Phi_n)) \to \phi(\omega)\nu(h)\int \Gamma_\theta(y)dy. \tag{C.17}$$

We have the following inversion formula

$$\tilde{\mathbb{E}}_\Phi^\kappa(\Gamma_\theta(y_n - \alpha n)h(\Phi_n)) = \frac{1}{2\pi}\int_{-M}^{M}\hat\Gamma_\theta(u)\tilde P_{\kappa,u}^n(h)du$$

where M is such that $\hat\Gamma_\theta(u) = 0$ outside $[-M, M]$.

Next, take a small ε_0. Then for $|u| < \varepsilon_0$ we have the decomposition

$$\tilde P_{\kappa,u}(h) = \lambda_{\kappa,u} m_{\kappa,u}(h) h_{\kappa,u} + \mathcal{R}_{\kappa,u}$$

where

$$\mathcal{R}_{\kappa,u}(h_{\kappa,u}) = 0, \quad m_{\kappa,u}(R_{\kappa,u}h) = 0 \text{ and } \|\mathcal{R}_{\kappa,u}^n\| \le K\bar\theta^n \text{ for some } \bar\theta < 1.$$

It follows that

$$\sqrt{n}\int_{|u|<\varepsilon_0}\hat\Gamma_\theta(u)\tilde P_{\kappa,u}^n(h)du = \frac{\sqrt{n}}{2\pi}\int_{-\varepsilon_0}^{\varepsilon_0}\hat\Gamma_\theta(u)\lambda_{\kappa,u}^n m_{\kappa,u}(h)\ h_{\kappa,u}du + \mathcal{O}(\bar\theta^n).$$

Next, letting $u = \frac{t}{\sqrt{n}}$ we can rewrite above integral as

$$\frac{1}{2\pi}\int_{-\varepsilon_0\sqrt{n}}^{\varepsilon_0\sqrt{n}}\Gamma(t/\sqrt{n})\lambda_{\kappa,t/\sqrt{n}}^n m_{\kappa,t/\sqrt{n}}(h)h_{\kappa,t/\sqrt{n}}dt. \tag{C.18}$$

The computations of the previous section give $\lambda_{\kappa,0} = 1$, $\lambda'_{\kappa,0} = 0$ so that $\lambda_{\kappa,t/\sqrt{n}}^n \to e^{-\frac{\sigma^2 t^2}{2}}$, where $\sigma^2 = \lambda''_\kappa$, and this convergence is dominated, that is, if ε_0 is small enough and $|t| \le \varepsilon_0\sqrt{n}$ we have $\lambda_{\kappa,t/\sqrt{n}}^n \le e^{-\frac{\sigma^2 t^2}{4}}$. As $u \to 0$ we have $m_{\kappa,u} \to m_\kappa$, $h_u \to 1$ so that the integral (C.18) converges to $\sigma\hat\Gamma_\theta(0)m_\kappa(h)$. On the other hand since for $\varepsilon_0 \le |u| \le M$ the spectral radius of $\tilde P_{\kappa,u}$ is strictly less than 1 we have

$$\frac{\sqrt{n}}{2\pi}\int_{\varepsilon_0<|u|<M}\hat\Gamma_\theta(u)\tilde P_{\kappa,u}^n(h)du = \mathcal{O}(\sqrt{n}\bar\theta^n)$$

and (C.17) follows.

Appendix D. Contracting property of the Markov chain ψ_n.

For two stochastic matrices ψ'_0 and ψ''_0 define ψ'_n and ψ''_n for $n \ge 1$ using the second formula in (2.1). The sequence $\{P_k, Q_k, R_k\}_{1\le k\le n}$ used in (2.1) is in both cases the same one.

Our goal is to estimate the norm $||\psi''_n - \psi'_n||$. We shall prove the following

Proposition D.1. *Assume that condition* (1.5) *holds. Then there are constants* $K = K(\varepsilon) > 0$, $\theta = \theta(\varepsilon) < 1$ *such that*

$$||\psi''_n - \psi'_n|| \le K\theta^n||\psi''_0 - \psi'_0||.$$

Proof. Consider two walkers $X'(t)$ and $X''(t)$ starting at the same site (n, i) in the layer L_n and moving on $[L_0, L_{n+1}]$ with reflecting boundary conditions ψ'_0 and ψ''_0 at L_0 respectively and absorbing boundary conditions at L_{n+1}.

We shall show that there exists a coupling between the walkers such that

$$\mathbb{P}(X'(\tilde{T}'_{n+1}) \neq X''(\tilde{T}''_{n+1})) \leq K\theta^n ||\psi''_0 - \psi'_0||, \tag{D.1}$$

where $\tilde{T}'_{n+1}$ and $\tilde{T}''_{n+1}$ are the hitting times of L_{n+1} for X' and X'' respectively. The statement then follows since, according to the probabilistic meaning of ψ'_n and ψ''_n,

$$\psi'_n(i,j) = \mathbb{P}(X'(\tilde{T}'_{n+1}) = j), \quad \psi''_n(i,j) = \mathbb{P}(X''(\tilde{T}''_{n+1}) = j)$$

and therefore

$$\sum_{j=1}^{m} |\psi'_n(i,j) - \psi''_n(i,j)| \leq 2\mathbb{P}(X'(\tilde{T}'_{n+1}) \neq X''(\tilde{T}''_{n+1})).$$

The coupling is constructed as follows.

1. The walkers walk together until they either reach L_0 for the first time or reach L_{n+1} without visiting L_0 (after that they stop). Note that the trajectories which miss L_0 do not contribute to the left hand side of (D.1).

2. Let t_0 be the first time the walkers reach L_0 and $t'_1 - 1$, $t''_1 - 1$ be the last time X' and X'' reach L_0 before reaching L_{n+1}. Between t_0 and t'_1 and between t_0 and t''_1 the walkers move independently.

Note that after time t_1 the walkers again move in the same environment $\{P_k, Q_k, R_k\}_{1\leq k\leq n}$ but conditioned on reaching L_{n+1} before L_0.

3. If $X'(t_1(X')) = X''(t_1(X''))$ then they move together until they reach L_{n+1}.

4. Denote by t'_l, t''_l the hitting times of L_l for X' and X'', where $l \geq 2$. If $X'(t_1(X')) \neq X''(t_1(X''))$ then they continue to move independently until they reach L_k with the minimal k such that $X'(t'_k) = X''(t''_k)$.

Note that the only trajectories that contribute to the left hand side of (D.1) are those for which $X'(t'_k) \neq X''(t''_k)$ for all $k \in [2, n+1]$.

We shall use the following estimate

Lemma D.2. *Let $X'(t'_1) = (1, J')$, $X''(t''_1) = (1, J'')$. There exists a constant C such that*

$$d(J', J'') \leq C||\psi''_0 - \psi'_0||$$

where d denotes the variational distance between the corresponding distributions.

If at step k of the procedure described above the walkers are uncoupled then condition (1.5) guarantees that the probability that they become coupled at L_{k+1} is at least εm. Thus the probability that the walkers were uncoupled at time t_1 and remain uncoupled until step $n + 1$ is less than

$$Cd(J', J'')\theta^n$$

and the result follows from Lemma D.2. □

Proof of Lemma D.2. Denote

$$B(j,k) = \mathbb{P}(X(t_1) = (1,k)|X(t_1+1) = (1,j)), \quad a(j) = \mathbb{P}(\tilde{T}_{n+1} < \tilde{T}_1|X(t_0+1) = (1,j)),$$

$$\Gamma(j,k) = \mathbf{P}(X \text{ returns to (1,k) after visiting } L_2 \text{ but before } \tilde{T}_{n+1}|X(\bar{t}+1) = (1,j)).$$

Also note that due to ellipticity we have

$$\min_j a_j \geq \varepsilon \max_j a_j.$$

Let M be the matrix with $M(j,k) = a(j)$.

With this notation our goal is to establish Lipshitz dependence of B on the boundary condition ψ. We have

$$B = M + Q\psi B + RB + \Gamma B$$

that is

$$B = (I - Q\psi - R - \Gamma)^{-1} M.$$

Therefore if B' and B'' correspond to different boundary conditions ψ' and ψ'' respectively then

$$B' - B'' = (I - Q\psi' - R - \Gamma)^{-1} Q(\psi' - \psi'')(I - Q\psi'' - R - \Gamma)^{-1} M.$$

The estimate we need is a consequence of two inequalities below.

$$||(\psi' - \psi'')(I - Q\psi'' - R - \Gamma)^{-1}|| \leq C||\psi' - \psi''|| \tag{D.2}$$

$$||(I - Q\psi' - R - \Gamma)^{-1}|| \leq C||M||^{-1} \tag{D.3}$$

To prove (D.2) let $U = Q\psi'' + R + \Gamma$. Given a probability vector p let $\pi'_k = p\psi' U^k$, $\pi''_k = p\psi'' U^k$. Due to ellipticity

$$||\pi'_k - \pi''_k|| \leq C(1-\varepsilon)^k ||\pi'_0 - \pi''_0|| \leq \tilde{C}(1-\varepsilon)^k ||\psi' - \psi''||$$

proving (D.2).

To prove (D.3) let $U = Q\psi' + R + \Gamma$. We have

$$\sum_k U(j,k) = 1 - a_j.$$

Hence

$$||U|| = \max_j \sum_k U(j,k) = 1 - \min_j a_j \leq 1 - \varepsilon \max_j a_j = 1 - \varepsilon||M||.$$

Accordingly $||U^k|| \leq (1 - \varepsilon||M||)^k$ so that $||(1-U)^{-1}|| \leq (\varepsilon||M||)^{-1}$ proving (D.3). □

References

[1] E. Bolthausen, I. Goldsheid: Recurrence and transience of random walks in random environments on a strip. Commun. Math. Phys. **214**, 429–447 (2000).

[2] E. Bolthausen, I. Goldsheid: Lingering random walks in random environment on a strip. Comm. Math. Phys. **278**, no. 1, 253–288 (2008).

[3] J. Brémont: On some random walks on $\mathbb{Z}$ in random medium, Ann. Prob. **30**, 1266–1312 (2002).

[4] J. Brémont: Random walks on $\mathbb{Z}$ in random medium and Lyapunov spectrum, Ann. Inst. H. Poincare Prob/Stat, **40**, 309-336 (2004).

[5] Y. Derriennic: Sur la récurrence des marches aléatoires unidimensionnelles en environement aléatoire, C. R. acad. Sci. Paris Sér. I Math.**329**, no. 1, 65–70 (1999).

[6] D. Dolgopyat and I. Goldsheid: Quenched limit theorems for nearest neighbour random walks in 1D random environment. Commun. Math. Phys **315**, 241–277 (2012).

[7] R. Durrett: Probability: theory and examples. 2d ed. Duxbury Press, Belmont, CA, 1996.

[8] N. Enriquez, C. Sabot, O. Zindy: Aging and quenched localization for one-dimensional random walks in random environment in the sub-ballistic regime. Bull. Soc. Math. France **137** (2009) 423-452.

[9] H. Furstenberg and H. Kesten: Products of random matrices. Ann. Math. Statist. **31**, 457–469 (1960).

[10] N. Gantert, Z. Shi: Many visits to a single site by a transient random walk in random environment. Stochastic Process. Appl. **99** (2002) 159–176.

[11] I. Goldhseid: Simple Transient Random Walks in One-dimensional Random Environment. Probability Theory and Related Fields, 139, Number 1, pages 41–64, 2007.

[12] I. Goldhseid: Linear and sub-linear growth and the CLT for hitting times of a random walk in random environment on a strip. Probab. Theory Related Fields **141** (2008) 471–511.

[13] A. O. Golosov: Localization of random walks in one-dimensional random environments. Comm. Math. Phys. **92** (1984), no. 4, 491–506.

[14] Y. Guivarc'h, E. Le Page: Simplicité de spectres de Lyapunov et propriété d'isolation spectrale pour une famille d'opérateurs sur l'espace projectif. Random walks and Geometry (Ed. V. Kaimanovitch), De Gruyter, 181-259 (2004).

[15] H. Kesten: Random Difference Equations and Renewal Theory for Products of Random Matrices. Acta Math. **131**, 207-248, (1973).

[16] H. Kesten: A renewal theorem for random walk in a random environment. Probability (Proc. Sympos. Pure Math., Vol. XXXI, Univ. Illinois, Urbana, Ill., 1976), pp. 67–77. Amer. Math. Soc., Providence, R.I., 1977.

[17] H. Kesten, M.V. Kozlov, and F. Spitzer: Limit law for random walk in a random environment. Composito Mathematica **30**, 145–168 (1975).

[18] E. Key: Recurrence and transience criteria for a random walk in random environment. Ann. Prob. **12**, 529-560 (1984).

[19] J. Kingman: Poisson processes. Oxford Studies in Probability **3** Oxford Univ. Press, New York 104 pp. (1993).
[20] S. Lalley: An extension of Kesten's renewal theorem for random walk in a random environment. Adv. in Appl. Math. **7** (1986) 80–100.
[21] A.V. Letchikov: A criterion for linear drift, and the central limit theorem for one-dimensional random walks in a random environment. Russian Acad. Sci. Sb. Math. **79**, 73-92 (1993).
[22] C. Liverani: Invariant measures and their properties. A functional analytic point of view. In *Dynamical systems. Part II,* 185–237, Pubbl. Cent. Ric. Mat. Ennio Giorgi, Scuola Norm. Sup., Pisa, 2003.
[23] E. Mayer-Wolf, A. Roitershtein, O. Zeitouni: Limit theorems for one-dimensional random walks in Markov random environments, Ann. Inst. H. Poincare Prob/Stat, **40**, 635-659 (2004).
[24] J. Peterson: Limiting distributions and large deviations for random walks in random environments. PhD Thesis– University of Minnesota (2008).
[25] J. Peterson: Quenched limits for transient, ballistic, sub-gaussian one-dimensional random walk in random environment. Ann. Inst. H. Poincare Probab. Statist. **45** 685-709 (2009).
[26] J. Peterson, O. Zeitouni: Quenched limits for transient, zero-speed one-dimensional random walk in random environment. Ann. Prob. **37** 143–188 (2009).
[27] A. Roitershtein: Tranzient random walks on a strip in a random environment. Ann. Probab. **36** 2354–2387 (2009).
[28] G. Samorodnitsky, M. S. Taqqu: Stable non-Gaussian random processes. Stochastic models with infinite variance. Stochastic Modeling. Chapman & Hall, New York 632 pp (1994).
[29] Ya. G. Sinai: The limiting behavior of a one-dimensional random walk in a random medium. Theory Prob. Appl. **27**, 256–268 (1982).
[30] Ya. G. Sinai: Simple random walk on tori. Journal of Stat. Physics, **94**, No 3-4, 695–708 (1999).
[31] F. Solomon: Random walks in a random environment. Ann. Prob. **3**, 1–31 (1975).
[32] S. Waddington: Large deviation asymptotics for Anosov flows. Ann. Inst. H. Poincare Anal. Nonlin. **13** (1996) 445–484.
[33] O. Zeitouni: Random walks in random environment, XXXI Summer school in Probability, St. Flour (2001). Lecture notes in Math. **1837**, 193-312, Springer, Berlin, 2004.

Dmitry Dolgopyat: Department of Mathematics and Institute of Physical Science and Technology, University of Maryland, College Park, MD, 20742, USA

Ilya Goldshied: School of Mathematical Sciences, Queen Mary University of London, London E1 4NS, Great Britain