ALEA, Lat. Am. J. Probab. Math. Stat. 8, 269–284 (2011)

On the range, local times and periodicity of random walk on an interval

Siva Athreya[1], Sunder Sethuraman[2] and Bálint Tóth[3]

Indian Statistical Institute
8th Mile Mysore Road, Bangalore 560059 India
and
Department of Mathematics, National University of Singapore
10 Lower Kent Ridge Road, 119076 Singapore.
E-mail address: athreya@isibang.ac.in
URL: http://www.isibang.ac.in/~athreya

Department of Mathematics, University of Arizona
617 N. Santa Rita Ave., Tucson, AZ 85721.
E-mail address: sethuram@math.arizona.edu
URL: http://math.arizona.edu/~sethuram

Institute of Mathematics, Technical University Budapest
Egry Jozsef u. 1
H-1111 Budapest, Hungary.
E-mail address: balint@math.bme.hu
URL: http://www.math.bme.hu/~balint

Abstract. The range, local times, and periodicity of symmetric, weakly asymmetric and asymmetric random walks at the time of exit from a strip with N locations are considered. Several results on asymptotic distributions are obtained.

1. Introduction

In this article, we study the range, local times, and periodicity or "parity" statistics of nearest-neighbor symmetric, weakly asymmetric, and asymmetric random walks up to the time of exit from an interval of N sites. We derive several associated scaling limits, as $N \to \infty$, which appear curious, some which connect with the entropy of an exit distribution, generalized Ray-Knight constructions, and Bessel and Ornstein-Uhlenbeck square processes, among other objects.

Received by the editors September 17, 2010; accepted August 24, 2011.
2000 *Mathematics Subject Classification.* primary 60K35; secondary 82C20.
Key words and phrases. range, random walk, local times, periodicity.
[1]Research supported in part by a CSIR Grant and Homi Bhaba Fellowship.
[2]Research supported in part by NSA-Q H982301010180, NSF-DMS 0906713.
[3]Research supported in part by a OTKA, grant no. K60708.

The study of the range of random walk is of course an old subject. However, examining the range and related statistics at the time the random walk leaves an interval, although a simple, natural concern, seems unexplored. We refer to Bass et al. (2009)[Ch. 2] and den Hollander and Weiss (1994) for exhaustive references on the range and related statistics of random walk in various settings.

From another view, indeed our initial motivation for this problem, the study of the range and other structures of random walk when it exits an interval can be thought of as a stochastic version of the "locker" problem, popular in university curriculum: Suppose there is a hallway of lockers labeled from 1 to N, for $N \geq 1$, which are initially closed. Let persons L, for $L \geq 1$, walk through the hallway, toggling every Lth locker, that is opening it if closed and closing it if open. The question is then to find out those lockers which will be open after the first N people walk through. The lockers whose labels are the squares, 1, 4, 9, etc., are exactly those with an odd number of factors. Consequently, these lockers are the open ones. Other variations of this problem can be found in Tanton (2007) and references therein.

In our random walk setting, we can imagine each site in the interval to be either open or closed, and the random walker toggling a site on each visit (from open to closed and vice versa) before it exits the interval. In comparison to the "locker" problem, we address the following questions:-

(1) What fraction of sites will be visited when the walker exits, e.g. the range?
(2) How many times will each site be visited before exit, e.g local times across sites?
(3) And, given a set of sites that have been visited, what is the joint distribution of their open status at the time of exit, e.g parity of the visits to points in the interval?

The specific answers naturally depend on the type of random walk considered. A goal of the paper is to see how the behaviors under symmetric and asymmetric walks are interpolated in terms of weakly asymmetric walks.

For the first question, we derive the limiting distribution for the range (Proposition 2.1), and observe as a consequence, which seems surprising, that the scaled range, when starting at random, is uniformly distributed on $[0,1]$ no matter the dynamics (Proposition 2.3). Also, curious values for the expected scaled range under symmetric walks, and the chance a given point is in the range, when starting at random are found (Remarks 2.2 and 2.4).

For the second question, we find the scaling limit of the local times through a "Ray-Knight" construction involving Bessel and Orstein-Uhlenbeck squared processes (Propositions 3.2, 3.4 and 3.5).

For the third question, we show that the parities of well-separated points, given that they are visited, are independent and identically distributed Bernoulli variables, and fair in the symmetric/weakly asymmetric case, and biased in the asymmetric situation (Proposition 4.1 and 4.2).

Set up: Let $\mathcal{T}_N = \{0, 1, 2, \ldots, N\}$. Let X_n be the position of a random walk on $\mathcal{T}_N$ at times $n \geq 1$. At each time step, the walk moves to the nearest point to its left (right) with probability q_N (p_N) where $p_N + q_N = 1$. The walk stops the moment it is at either 0 or N. When $p_N = q_N = 1/2$, the walk is of course referred to as the symmetric random walk. When $q_N = 1/2 - c/N$ (and so $p_N = 1/2 + c/N$)

for some constant $c > 0$ and N large enough so that $0 < p_N, q_N < 1$, we say the walk is weakly asymmetric. When $q_N = q < p = p_N$, the walk is asymmetric.

Define $T_a = \inf\{n \geq 1 : X_n = a\}$ as the hitting time of $a \in \mathcal{T}_N$. Then, $\tau_N = T_0 \wedge T_N$ is the "exit" time from the strip. Clearly, starting from $1 \leq x \leq N-1$, τ_N is finite: $P_x(\tau_N < \infty) = 1$ where we denote $P_x(A) = P(A|X_0 = x)$ as the conditional probability of the event A with respect to the walk starting from $X_0 = x$.

Then, the number of visits to $y \in \mathcal{T}_N$ before exiting is $G(y) = \sum_{k=0}^{\tau_N} 1_y(X_k)$. Hence, the event y is visited at all corresponds to $G(y) \geq 1$. In this case, we say the parity of y is "even" (locker y is closed) if $G(y) \geq 1$ and $G(y) = 0 \bmod_2$. Correspondingly, the parity of y is "odd" (locker y is open) when $G(y) \geq 1$ and $G(y) = 1 \bmod_2$.

The plan of the article is to address questions (1), (2) and (3) in sections 2, 3, and 4 respectively.

2. Question 1: Range of random walk in $\mathcal{T}_N$

In this section, we obtain distributional limits of the range up to the exit time when starting from a point, and at random in subsections 2.1, 2.2.

2.1. *The range starting from a point.* Denote R_N as the number of locations visited before exit, the range of the walk on $\mathcal{T}_N$, that is

$$R_N = \#\{y \in \mathcal{T}_N : G(y) \geq 1\}.$$

Observe, when starting from $[\alpha N]$, necessarily $[\alpha N \wedge (1-\alpha)N] \leq R_N \leq N$.

Proposition 2.1. *Let $X_0 = [\alpha N]$ for $0 < \alpha < 1$. For symmetric and weakly asymmetric walks, R_N/N converges in distribution to absolutely continuous measures on $[0,1]$, respectively $F_{0,\alpha}$ and $F_{c,\alpha}$ defined in (2.2) and (2.5). For asymmetric walks, $R_N - [(1-\alpha)N] \Rightarrow Z$ where Z is Geometric(q/p).*

Proof: First, we observe, for $0 < \beta < 1$, starting from location $x = [\alpha N]$, since the motion is nearest-neighbor,

$$\begin{aligned}
\{R_N \geq \beta N\} &= \{R_N \geq \beta N, \tau_N = T_0\} \cup \{R_N \geq \beta N, \tau_N = T_N\} \\
&= \{T_{[\beta N]} < T_0 < T_N\} \cup \{T_{N-[\beta N]} < T_N < T_0\}.
\end{aligned}$$

We now specialize arguments to the three types of random walks.

Symmetric Walk:- When the walk is symmetric $p_N = q_N = 1/2$, recall the standard Gambler's ruin identity: For $a, b, z \in \mathcal{T}_N$, such that $a < z < b$,

$$P_z(T_a < T_b) = \frac{b-z}{b-a}. \tag{2.1}$$

For $\beta > \alpha$, compute

$$P_{[\alpha N]}(R_N \geq [\beta N], \tau_N = T_0) = \frac{[\alpha N]}{[\beta N]}\frac{N-[\beta N]}{N} \to \frac{\alpha(1-\beta)}{\beta}.$$

When, $\beta > 1-\alpha$, we have

$$P_{[\alpha N]}(R_N \geq [\beta N], \tau_N = T_N) = \frac{N-[\alpha N]}{[\beta N]}\frac{N-[\beta N]}{N} \to \frac{(1-\alpha)(1-\beta)}{\beta}.$$

Putting these expressions together, along with simple calculations, we have

$$\lim_{N\uparrow\infty} P_{[\alpha N]}(R_N/N \geq \beta) \;=\; \begin{cases} 1 & \text{when } 0 \leq \beta \leq \alpha\wedge(1-\alpha) \\ \frac{\alpha\wedge(1-\alpha)}{\beta} & \text{when } \alpha\wedge(1-\alpha) < \beta < \alpha\vee(1-\alpha) \\ \frac{1-\beta}{\beta} & \text{when } \alpha\vee(1-\alpha) \leq \beta \leq 1 \\ 0 & \text{when } \beta > 1. \end{cases}$$

The right-side defines a distribution $F_{0,\alpha}$, supported on $[\alpha\wedge(1-\alpha), 1]$ whose density

$$f_\alpha(\beta) \;=\; \begin{cases} \frac{\alpha\wedge(1-\alpha)}{\beta^2} & \text{for } \alpha\wedge(1-\alpha) < \beta < \alpha\vee(1-\alpha) \\ \frac{1}{\beta^2} & \text{for } \alpha\vee(1-\alpha) \leq \beta \leq 1 \\ 0 & \text{otherwise.} \end{cases} \tag{2.2}$$

Weakly-asymmetric walk:- In the weakly asymmetric case, $q_N = 1/2 - c/N$ and $p_N = 1/2 + c/N$ with $c > 0$, let

$$s_N \;:=\; \frac{q_N}{p_N} \;=\; \frac{1/2 - c/N}{1/2 + c/N} \;=\; 1 - \frac{4c}{N} + O(N^{-2}). \tag{2.3}$$

The corresponding gambler's ruin identity becomes, for $a < z < b$,

$$P_z(T_a < T_b) \;=\; \frac{(q_N/p_N)^z - (q_N/p_N)^b}{(q_N/p_N)^a - (q_N/p_N)^b}. \tag{2.4}$$

Then, following the symmetric argument, when $\beta > \alpha$,

$$\begin{aligned} P_{[\alpha N]}(R_N \geq [\beta N], T_0 = \tau_N) \;&=\; \frac{1 - s_N^{[\alpha N]}}{1 - s_N^{[\beta N]}} \frac{s_N^{[\beta N]} - s_N^N}{1 - s_N^N} \\ &\to\; \frac{1 - e^{-4\alpha c}}{1 - e^{-4\beta c}} \frac{e^{-4\beta c} - e^{-4c}}{1 - e^{-4c}} \;:=\; A_1(\alpha,\beta,c). \end{aligned}$$

When $\beta > 1 - \alpha$,

$$\begin{aligned} P_{[\alpha N]}(R_N \geq [\beta N], T_N = \tau_N) \;&=\; \frac{s_N^{[\alpha N]} - s_N^N}{s_N^{N-[\beta N]} - s_N^N} \frac{1 - s_N^{N-[\beta N]}}{1 - s_N^N} \\ &\to\; \frac{e^{-4\alpha c} - e^{-4c}}{e^{-4(1-\beta)c} - e^{-4c}} \frac{1 - e^{-4(1-\beta)c}}{1 - e^{-4c}} \;:=\; A_2(\alpha,\beta,c). \end{aligned}$$

Noting

$$\lim_{N\uparrow\infty} P_{[\alpha N]}(T_N < T_0) \;=\; \frac{1 - e^{-4\alpha c}}{1 - e^{-4c}}, \quad \text{and} \quad A_1 + A_2 \;=\; \frac{e^{-4\alpha c}(e^{4c(1-\beta)} - 1)}{1 - e^{-4c\beta}},$$

we have

$$\lim_{N\uparrow\infty} P_{[\alpha N]}(R_N/N \geq \beta) \;=\; \begin{cases} 1 & \text{when } 0 \leq \beta \leq \alpha\wedge(1-\alpha) \\ \frac{1-e^{-4c\alpha}}{1-e^{-4c\beta}} & \text{when } \alpha \leq \beta \leq 1-\alpha \\ \frac{1-e^{4c(1-\alpha)}}{1-e^{4c\beta}} & \text{when } 1-\alpha \leq \beta \leq \alpha \\ \frac{e^{-4\alpha c}(e^{4c(1-\beta)}-1)}{1-e^{-4c\beta}} & \text{when } \alpha\vee(1-\alpha) \leq \beta \leq 1 \\ 0 & \text{when } \beta > 1. \end{cases} \tag{2.5}$$

The right-side defines a distribution $F_{c,\alpha}$, supported on $[\alpha\wedge(1-\alpha), 1]$, whose density, although messy, can be easily derived.

Asymmetric walk:- In the asymmetric case, $q_N = q$, $p_N = p$, and $p > q$, and it is not difficult to see that we cannot go left too many times. The gambler's ruin identity (2.4) also holds in this case and, for $x = [\alpha N]$, $P_x(T_0 < T_N) = \exp\{-CN\}$ for some constant $C > 0$.

To complete the proof, for integers $z \geq 0$, compute

$$\begin{aligned} P_x(R_N \geq N - x + z) &= P_x(T_N < T_0, R_N \geq N - x + z) + o(1) \\ &= P_x(T_{x-z} < T_N < T_0) + o(1) \\ &= (q/p)^z + o(1). \end{aligned}$$

□

Remark 2.2. 1. As expected, $F_{c,\alpha}$ interpolates between the symmetric and asymmetric cases: Namely, as $c \downarrow 0$, $F_{c,\alpha} \Rightarrow F_{0,\alpha}$, and as $c \uparrow \infty$, $F_{c,\alpha}$ converges to the constant $1 - \alpha$.

2. It is curious to observe, for symmetric walks, starting from $x = [\alpha N]$, with $\alpha \in (0, 1/2]$, the expected range

$$\int_0^1 \beta f_\alpha(\beta) d\beta = \int_\alpha^{1-\alpha} \beta \frac{\alpha}{\beta^2} d\beta + \int_{1-\alpha}^1 \beta \frac{1}{\beta^2} d\beta = -(1-\alpha)\log(1-\alpha) - \alpha \log(\alpha)$$

is the entropy of the exit distribution $\langle 1 - \alpha, \alpha \rangle$ where $1 - \alpha$ is the probability of leaving by the left endpoint, and α the chance of exiting right! The maximum value $\log 2$ occurs when $\alpha = 1/2$.

3. For symmetric and weakly asymmetric walks, the limit distributions may also be derived in terms of Brownian motion and diffusion estimates.

2.2. *The range when starting at random.* We derive now the limiting law of R_N/N when starting at random, that is the uniform distribution on $\mathcal{T}_N$. It seems nonintuitive that the limit law is U$[0, 1]$ no matter the type of random walk.

Proposition 2.3. *For symmetric, weakly asymmetric and asymmetric random walk, when starting at random in $\mathcal{T}_N$, R_N/N converges weakly to the uniform distribution $U[0, 1]$.*

Proof: Suppose our starting point was random. In the symmetric and weakly asymmetric cases, the limiting distribution of R_N/N, from straightforward considerations, is found by integrating the density f_α and tail of $F_{c,\alpha}$ with respect to α (denoted by $F_{c,\alpha}([\beta, 1])$).

In the symmetric case, when $\beta \leq 1/2$,

$$\int_0^1 f_\alpha(\beta) d\alpha = \int_0^\beta \frac{\alpha}{\beta^2} d\alpha + \int_\beta^{1-\beta} 0 \, d\alpha + \int_{1-\beta}^1 \frac{1-\alpha}{\beta^2} d\alpha = 1.$$

But, also, when $\beta > 1/2$,

$$\int_0^1 f_\alpha(\beta) d\alpha = \int_0^{1-\beta} \frac{\alpha}{\beta^2} d\alpha + \int_\beta^{1-\beta} \frac{1}{\beta^2} d\alpha + \int_\beta^1 \frac{1-\alpha}{\beta^2} d\alpha = 1.$$

On the other hand, in the weakly asymmetric case, we have, when $\beta \leq 1/2$,

$$\begin{aligned} \int_0^1 F_{c,\alpha}([\beta, 1]) d\alpha &= \int_0^\beta \frac{1 - e^{-4c\alpha}}{1 - e^{-4c\beta}} d\alpha + \int_\beta^{1-\beta} 1 \, d\alpha \\ &\quad + \int_{1-\beta}^1 \frac{e^{-4c\beta}(e^{4c(1-\alpha)} - 1)}{e^{-4c(1-\beta)} - e^{-4c}} d\alpha = 1 - \beta. \end{aligned}$$

Similarly, when $\beta > 1/2$, $\int_0^1 F_{c,\alpha}([\beta,1])d\alpha$ equals

$$\frac{1}{1-e^{-4c\beta}}\Big[\int_0^{1-\beta} 1-e^{-4c\alpha}d\alpha + \int_{1-\beta}^{\beta} e^{-4c\alpha}(e^{4c(1-\beta)}-1)d\alpha + \int_\beta^1 e^{-4c\beta}(e^{4c(1-\alpha)}-1)d\alpha\Big] \ = \ 1-\beta.$$

Consequently, for symmetric and weakly asymmetric walks, the limiting distribution is $U[0,1]$ when the starting position is uniformly chosen.

However, in the asymmetric case, from Proposition 2.1, $R_N/N \to 1-\alpha$ in probability starting from $x = [\alpha N]$. Then, starting at random in $\mathcal{T}_N$, we have that $R_N/N \to Y$ in probability where Y is a $U[0,1]$ distributed random variable. □

Remark 2.4. One might ask, on the other hand, with what probability a point $y = [\beta N]$ belongs to the range when starting at random. This is the same as asking when y is visited by the walk. For symmetric walk, it is not difficult to use the gambler's ruin identity (2.1) to see, as $N \uparrow \infty$, that the probability tends to

$$\int_\beta^1 \frac{1-\alpha}{1-\beta}d\alpha + \int_0^\beta \frac{\alpha}{\beta}d\alpha \ = \ \frac{1}{2}.$$

It seems curious that the limit does not depend on β.

For asymmetric walk, starting from $[\alpha N]$, when $\alpha > \beta$, the point y cannot be reached with positive probability in the limit. Then, the chance y belongs to the range, when starting at random, is β.

For weakly asymmetric walks, using (2.4), the limit is $\frac{\beta}{1-e^{-4c\beta}} - \frac{1-\beta}{1-e^{4c(1-\beta)}}$ which interpolates between the other cases as $c \downarrow 0$ and $c \uparrow \infty$

3. Question 2: Characterization of local times

To capture the local times of the random walk before its exit, we use the "Ray-Knight" or "Kesten-Kozlov-Spitzer" representation, and some martingale characterizations. Our treatment and proofs will be similar to those in Tóth (1996) which considered certain self-interacting random walks.

Let $0 < \alpha < 1$. Suppose the walk starts at $[\alpha N]$, and exits at the right endpoint N. Let ζ_j^N be the number of left crossings of the bond $(N-j-1, N-j)$ before exit. Then, $\zeta_0^N = 0$, and ζ_1^N is distributed as D_N, a Geometric(q_N) random variable minus 1, $P(D_N = n) = p_N q_N^n$ for $n \geq 0$. In the following, we drop the script N.

Let $\{\xi_{j,i}\}_{i,j\geq 0}$ be i.i.d. random variables with distribution D_N. A moment's thought convinces that $\{\zeta_j\}_{0\leq j\leq N}$ is a Markov chain with representation

$$\zeta_{j+1} \ = \ \begin{cases} \sum_{i=0}^{\zeta_j} \xi_{j,i} & \text{for } 0 \leq j < [(1-\alpha)N] \\ \sum_{i=1}^{\zeta_j} \xi_{j,i} & \text{for } [(1-\alpha)N] \leq j \leq N-1 \end{cases} \tag{3.1}$$

such that

$$\zeta_j = 0 \quad \text{for some} \quad [(1-\alpha)N] \leq j < N, \tag{3.2}$$

with the convention that empty sums vanish.

Note that for $j < [(1-\alpha)N]$, the sum starts with index $i = 0$ since, even if $\zeta_j = 0$, given exit at the right, the walk must visit locations $[\alpha N] \leq x \leq N$ and may have left crossings of $(x-1, x)$. However, for $j \geq [(1-\alpha)N]$, since the walk is not guaranteed to visit sites to the left of $[\alpha N]$, ζ_j is the size of a branching process, with initial value $\zeta_{[(1-\alpha)N]}$, which must vanish before time $j = N$.

Then, the local time of the walk is

$$G(y) = \begin{cases} \zeta_{N-y} & \text{for } 0 \le y < [\alpha N] \\ \zeta_{N-y} + 1 & \text{for } [\alpha N] \le y \le N. \end{cases}$$

In the following, to analyze $\{\zeta_j\}_{0\le j\le N}$, it will be helpful to consider the Markov chain η_j, such that $\eta_0 = 0$ and $\eta_1 \stackrel{d}{=} D_N$, for which representation (3.1) holds in terms of the variables $\{\xi_{j,i}\}_{i,j\ge 0}$, but *without* the restriction (3.2).

When the walk exits at the left endpoint 0, one considers an analogous Markov chain $\tilde{\zeta}_j$, corresponding to right-crossings of $(j, j+1)$, where the representation and restriction are reversed. Namely, let $\tilde{D}_N$ be a Geometric(p_N) random variable minus 1, $P(\tilde{D}_N = n) = q_N p_N^n$ for $n \ge 0$. Define $\tilde{\zeta}_0 = 0$, $\tilde{\zeta}_1 \stackrel{d}{=} \tilde{D}_N$, and

$$\zeta_{j+1} = \begin{cases} \sum_{i=0}^{\zeta_j} \xi_{j,i} & \text{for } 0 \le j < [\alpha N] \\ \sum_{i=1}^{\zeta_j} \xi_{j,i} & \text{for } [\alpha N] \le j \le N-1 \end{cases}$$

such that $\zeta_j = 0$ for some $[\alpha N] \le j < N$. The local time of the walk in this case is $G(y) = \tilde{\zeta}_y$ for $[\alpha N] < y \le N$ and $G(y) = \tilde{\zeta}_y + 1$ for $0 \le y \le [\alpha N]$. Here also it will be of use to define analogously a Markov chain $\tilde{\eta}_j$ satisfying $\tilde{\eta}_0 = 0$, $\tilde{\eta}_1 \stackrel{d}{=} \tilde{D}_N$, and the reversed representation but without the restriction that the chain must vanish for $[\alpha N] \le j < N$.

Finally, define $Y_N(t) = \frac{1}{N}\eta_{[Nt]}$ and $\tilde{Y}_N(t) = \frac{1}{N}\tilde{\eta}_{[Nt]}$ for $0 \le t \le 1$, and suppose that $Y_N(0) = \tilde{Y}_N(0) = 0$.

3.1. *Symmetric walks.* Consider the following processes. Let $Z_0 = 0$, and define Z_t as a solution of the stochastic differential equation given by

$$Z_t = \begin{cases} t + \int_0^t \sqrt{2Z_s}\,dB_s & \text{for } 0 \le t \le 1-\alpha \\ \\ Z_{1-\alpha} + \int_{1-\alpha}^t \sqrt{2Z_s}\,dB_s & \text{for } 1-\alpha \le t \le 1. \end{cases}$$

Observe that $\{Z_t\}_{0\le t\le 1-\alpha}$ is the same in law as $\{\mathrm{Besq}^2(t/2)\}_{0\le t\le 1-\alpha}$ process, and $\{Z_t\}_{0\le t\le 1-\alpha}$ is the same in law as $\{\mathrm{Besq}^0(t/2)\}_{1-\alpha\le t\le 1}$ process. For more on the processes $\{\mathrm{Besq}^\delta(t)\}_{t\ge 0}$ i.e. solutions to $dX_t = \delta dt + 2\sqrt{X_t}dB_t$ please see Revuz and Yor (1999) . Let τ_0^R be the first time Z_t hits 0 after time $t = 1-\alpha$. Note that Z_t remains at value 0 after time τ_0^R.

Define also $\tilde{Z}_0 = 0$ and define $\tilde{Z}_t$ as a solution of the stochastic differential equation given by

$$\tilde{Z}_t = \begin{cases} t + \int_0^t \sqrt{2\tilde{Z}_s}\,dB_s & \text{for } 0 \le t \le \alpha \\ \tilde{Z}_\alpha + \int_\alpha^t \sqrt{2\tilde{Z}_s}\,dB_s & \text{for } \alpha \le t \le 1. \end{cases}$$

Let also τ_0^L be the time $\tilde{Z}_t$ reaches 0 after time $t = \alpha$. Here, also, $\tilde{Z}_t \equiv 0$ for $t \ge \tau_0^L$.

It will turn out that Z_t and $\tilde{Z}_t$ will be identified respectively, as the scaling limits of the local times when the random walk exits at the right and left endpoints of the interval. The important point in this identification is the next result.

Proposition 3.1. *For symmetric walk starting from $x = [\alpha N]$, we have*

$$\{Y_N(t)\}_{0\le t\le 1} \Rightarrow \{Z_t\}_{0\le t\le 1} \quad \text{and} \quad \{\tilde{Y}_N(t)\}_{0\le t\le 1} \Rightarrow \{\tilde{Z}_t\}_{0\le t\le 1}$$

in $D[0,1]$, *in the sup topology.*

A similar result is established in Knight (1963). Instead of proving Proposition 3.1, which can also be done following steps in Tóth (1996) , we prove Proposition 3.3 in the next subsection, with respect to weakly asymmetric random walks, dealing with squared Ornstein-Uhlenbeck processes which are less standard.

Now, with Proposition 3.1 in hand, since Y_N and $\tilde{Y}_N$ converge respectively to Z and $\tilde{Z}$ in the sup topology, it follows that the conditional distributions of Y_N given that η_j vanishes for $j \geq [(1-\alpha)N]$ and $\tilde{Y}_N$ given that $\tilde{\eta}_j$ vanishes for $j \geq [\alpha N]$ converge to the conditional distributions of Z given that $1-\alpha \leq \tau_0^R < 1$ and $\tilde{Z}$ given that $\alpha \leq \tau_0^L < 1$.

Hence, from this discussion, the following characterization holds for the local times of the walk up to time of exit. Recall that $1-\alpha$ and α are the exit probabilities of right and left exit respectively.

Proposition 3.2. *For symmetric walk starting from* $[\alpha N]$, *the local times*

$$\{G([Nt])/N\}_{0\leq t\leq 1} \;\Rightarrow\; \alpha\mu^R + (1-\alpha)\mu^L$$

where μ^R *is the law of the process* $\{Z_{1-t}\}_{0\leq t\leq 1}$ *conditioned on* $1-\alpha \leq \tau_0^R < 1$, *and* μ^L *is the law of the process* $\{\tilde{Z}_t\}_{0\leq t\leq 1}$ *conditioned on* $\alpha \leq \tau_0^L < 1$.

3.2. *Weakly asymmetric walks.* The development of the local time structure is similar to the symmetric case. Corresponding to right exit, $ED_N = q_N/p_N = 1 - \frac{4c}{N+2c}$ and $\mathrm{Var}(D_N) = q_N^2/p_N^2 + q_N/p_N = 2 - \frac{12c}{N+2c} + \frac{16c^2}{(N+2c)^2}$. Let $Z_0^c = 0$ and define Z_t^c as a solution of the stochastic differential equation given by

$$Z_t^c \;=\; \begin{cases} \int_0^t (1-4cZ_s^c)ds + \int_0^t \sqrt{2Z_s^c}dB_s & \text{for } 0\leq t\leq 1-\alpha \\ \\ Z_{1-\alpha}^c - \int_{1-\alpha}^t 4cZ_s^c ds + \int_{1-\alpha}^t \sqrt{2Z_s^c}dB_s & \text{for } 1-\alpha \leq t\leq 1 \end{cases}$$

Note $2(Z_t^c+1)$ and $2(Z_t^c-t)$ are the squares of the Ornstein-Uhlenbeck process $dX_t = -4cX_t dt + \sqrt{2}dB_t$ for $0\leq t\leq 1-\alpha$ and $1-\alpha\leq t\leq 1$ respectively.

Also, with respect to left exit, $E\tilde{D}_N = p_N/q_N = 1 + 4c/N + O(N^{-2})$ and $\mathrm{Var}(\tilde{D}_N) = 2 + O(N^{-1})$. Define $\tilde{Z}_0^c = 0$ and define $\tilde{Z}_t^c$ as a solution of the stochastic differential equation given by

$$\tilde{Z}_t^c \;=\; \begin{cases} \int_0^t (1+4c\tilde{Z}_s^c)ds + \int_0^t \sqrt{2\tilde{Z}_s^c}dB_s & \text{for } 0\leq t\leq \alpha \\ \tilde{Z}_\alpha^c + \int_\alpha^t \sqrt{2\tilde{Z}_s^c}dB_s & \text{for } \alpha\leq t\leq 1. \end{cases}$$

As before, let $\hat{\tau}_0^R$ be the first time after $t = 1-\alpha$ that Z_t^c reaches 0, and $\hat{\tau}_0^L$ be the first time after $t=\alpha$ that $\tilde{Z}_t^c$ hits 0.

Analogous to the symmetric random walk case, we show that Z^c and $\tilde{Z}^c$ are the scaling limits of the local times when the weakly asymmetric random walk exits at the right and left endpoints respectively.

Proposition 3.3. *For the weakly asymmetric random walk starting from* $x = [\alpha N]$, *we have*

$$\{Y_N(t)\}_{0\leq t\leq 1} \;\Rightarrow\; \{Z_t^c\}_{0\leq t\leq 1} \quad\text{and}\quad \{\tilde{Y}_N(t)\}_{0\leq t\leq 1} \;\Rightarrow\; \{\tilde{Z}^c(t)\}_{0\leq t\leq 1}$$

in $D[0,1]$, *in the sup topology.*

The same argument as in the symmetric case allows to deduce the the following characterization.

Proposition 3.4. *For the weakly asymmetric walk, starting from $x = [\alpha N]$, the local times satisfy*

$$\{G([Nt])/N\}_{0\le t\le 1} \;\Rightarrow\; R(\alpha)\mu_c^R + (1-R(\alpha))\mu_c^L$$

where μ_c^R is the law of the process $\{Z_{1-t}^c\}_{0\le t\le 1}$ conditioned on $1-\alpha \le \hat{\tau}_0^R < 1$, and μ_c^L is the law of the process $\{\tilde{Z}_t^c\}_{0\le t\le 1}$ conditioned on $\alpha \le \hat{\tau}_0^L < 1$. Here, $R(\alpha) = (1-e^{-4c\alpha})/(1-e^{-4c})$ is the exit probability to the right.

Proof of Proposition 3.3. Here, we argue that $\{Y_N(t)\}_{0\le t\le 1} \Rightarrow \{Z_t^c\}_{0\le t\le 1}$ which corresponds to "left crossings." The argument for $\{\tilde{Y}_N(t)\}_{0\le t\le 1} \Rightarrow \{\tilde{Z}_t^c\}_{0\le t\le 1}$ is similar.

The proof naturally separates into two parts corresponding to when $j \le [(1-\alpha)N]$ and $j \ge [(1-\alpha)N]$. The strategy will be to use martingale decompositions of the Markov chain $\{\eta_j\}_{j\ge 0}$. Define, for $0\le t\le 1$ and $[Nt] \le [(1-\alpha)N]$, the martingale and its quadratic variation,

$$\begin{aligned}
M_N(t) &= \eta_{[Nt]} - \eta_0 - \sum_{j=0}^{[Nt]-1} \left(E[\eta_{j+1}|\eta_j] - \eta_j\right) \\
\langle M_N(t)\rangle &= \sum_{j=0}^{[Nt]-1} E\big[\big(\eta_{j+1} - E[\eta_{j+1}|\eta_j]\big)^2\big].
\end{aligned}$$

Then,

$$\begin{aligned}
\frac{1}{N}M_N(t) &= Y_N(t) - Y_N(0) - \frac{1}{N}\sum_{j=0}^{[Nt]-1}\left(E(D_N)(\eta_j+1) - \eta_j\right) \\
&= Y_N(t) - Y_N(0) - \frac{1}{N}\sum_{j=0}^{[Nt]-1}\left((1-\frac{4c}{N+2c})(\eta_j+1) - \eta_j\right) \\
&= Y_N(t) - Y_N(0) - \frac{1}{N}[Nt] + \frac{4c}{N+2c}\sum_{j=0}^{[Nt]-1} Y_N\left(\frac{j}{N}\right) + \frac{4c[Nt]}{N(N+2c)}
\end{aligned}$$

and

$$\begin{aligned}
\langle N^{-1}M_N(t)\rangle &= \frac{1}{N^2}\sum_{j=0}^{[Nt]-1} E\big[(\eta_{j+1} - E(D_N)(\eta_j+1))^2|\eta_j\big] \\
&= \frac{1}{N^2}\sum_{j=0}^{[Nt]-1} E\left[\left(\sum_{i=0}^{\eta_j}(\xi_{j,i} - E(D_N))\right)^2 |\eta_j\right] \qquad (3.3)\\
&= \frac{1}{N^2}\sum_{j=0}^{[Nt]-1}(\eta_j+1)\mathrm{Var}(D_N) \;=\; \frac{2}{N}\sum_{j=0}^{[Nt]-1} Y_N\left(\frac{j}{N}\right) + O\left(\frac{1}{N}\right).
\end{aligned}$$

Now suppose the two sequences in N, $\{Y_N(t)\}_{0\le t\le 1}$ and $\{N^{-1}M_N(t)\}_{0\le t\le 1}$, are tight in the sup topology, and $\{Y_N(t)\}_{0\le t\le 1} \Rightarrow \{Z_t\}_{0\le t\le 1}$, $\{N^{-1}M_N(t)\}_{0\le t\le 1} \Rightarrow \{M(t)\}_{0\le t\le 1}$ on subsequences. Then, for $0\le t\le 1-\alpha$, $M(t) = Z_t - Z_0 - \int_0^t (1-$

$4cZ_s)ds$ and $\langle M(t)\rangle = 2\int_0^t Z_s ds$. Hence, for $0 \le t \le 1-\alpha$, by Levy's criteria for continuous martingales, we have that Z_t is uniquely characterized by

$$Z_t = Z_0 + \int_0^t (1-4cZ_s)ds + \int_0^t \sqrt{2Z_s}dB_s$$

Similarly, for $0 \le t \le 1$ such that $[Nt] \ge [(1-\alpha)N]$, since now $\eta_{j+1} = \sum_{i=1}^{\eta_j} \xi_{j,i}$, the drift is not present, and we can write

$$\begin{aligned}\frac{1}{N}(M_N(t) - M_N(1-\alpha)) &= Y_N(t) - Y_N(1-\alpha) + \sum_{j=[(1-\alpha)N]}^{[Nt]-1} (E(\eta_{j+1}|\eta_j) - \eta_j) \\ &= Y_N(t) - Y_N(1-\alpha) + \frac{4c}{N+c}\sum_{j=[N(1-\alpha)]}^{[Nt]-1} Y_N\left(\frac{j}{N}\right)\end{aligned}$$

and also

$$\begin{aligned}\frac{1}{N}(\langle M_N(t)\rangle - \langle M_N(1-\alpha)\rangle) &= \frac{1}{N^2}\sum_{j=[N(1-\alpha)]}^{[Nt]-1} (\eta_j+1)\mathrm{Var}(D_N) \\ &= \frac{2}{N}\sum_{j=[N(1-\alpha)]}^{[Nt]-1} Y_N\left(\frac{j}{N}\right) + O\left(\frac{1}{N}\right).\end{aligned}$$

Hence, as before, given tightness of the sequence in N, $\{N^{-1}M_N(t)\}_{0\le t\le 1}$, and subsequential convergences $\{Y_N(t)\}_{0\le t\le 1} \Rightarrow \{Z_t\}_{0\le t\le 1}$ and $\{N^{-1}M_N(t)\}_{0\le t\le 1} \Rightarrow \{M(t)\}_{0\le t\le 1}$ where on $[1-\alpha, 1]$ $M(t) - M(1-\alpha) = Z_t - Z_{1-\alpha}$ and $\langle M(t) - M(1-\alpha)\rangle = 2\int_{1-\alpha}^t Z_s ds$, we conclude, for $t \in [1-\alpha, 1]$, that

$$Z_t = Z_{1-\alpha} + \int_{1-\alpha}^t \sqrt{2Z_s}dB_s.$$

Consequently, by putting the subsequential converges together, that $\{Y_N(t)\}_{0\le t\le 1}$ converges weakly to $\{Z_t\}_{0\le t\le 1}$.

Tightness. We now argue tightness of the sequences in N, $Y_N(t)$ and $N^{-1}M_N(t)$, for $t \in [0, 1-\alpha]$. Tightness of $\{Y_N(t)\}_{0\le t\le 1-\alpha}$ follows from tightness of $\{N^{-1}M_N(t)\}_{0\le t\le 1-\alpha}$ in the sup topology which can be argued by a Kolmogorov-Centsóv argument. First, for a general discrete time martingale $(M(l), \mathcal{F}_l)$ with difference $\delta(l) = M(l) - M(l-1)$, we have that

$$\begin{aligned}E\left[(M(l) - M(k))^4\right] &= 6\sum_{j=k+1}^{l} E\left[\delta(j)^2(M(j-1) - M(k))^2\right] \\ &\quad +4\sum_{j=k+1}^{l} E\left[\delta(j)^3(M(j-1) - M(k))\right] + \sum_{j=k+1}^{l} E\left[\delta_j^4\right]\end{aligned}$$

and by Jensen inequality,

$$E\left[(M(l)-M(k))^4\right] \le 6\sum_{j=k+1}^{l} E\left[E[\delta(j)^2|\mathcal{F}_{j-1}](M(j-1)-M(k))^2\right]$$
$$+\,4\sum_{j=k+1}^{l}\left\{E\left[E[\delta(j)^4|\mathcal{F}_{j-1}]^{3/2}(M(j-1)-M(k))^2\right]\right\}^{1/2}$$
$$+\sum_{j=k+1}^{l} E\left[E[\delta(j)^4|\mathcal{F}_{j-1}]\right].$$

Now, in our context, define the martingale, for $l \le [(1-\alpha)N]$,

$$M(l) = \eta_l - \eta_0 - \sum_{i=0}^{l-1}(E[\eta_{i+1}|\eta_i]-\eta_i)$$

so that $M_N(t) = M([Nt])$, and also the stopping time

$$\theta_{y,N} = \inf\{l \ge 0 : \eta_l \ge Ny\}.$$

Compute, with respect to $M(l\wedge\theta_{y,N})$, that

$$\begin{aligned}\delta(l) &= M(l\wedge\theta_{y,N}) - M(l-1\wedge\theta_{y,N})\\ &= \eta_{l\wedge\theta_{N,y}} - E[\eta_{l\wedge\theta_{y,N}}|\eta_{l-1\wedge\theta_{y,N}}] = \sum_{i=0}^{\eta_{l-1\wedge\theta_{N,y}}}(\xi_{l\wedge\theta_{y,N},i}-E(D_N)).\end{aligned}$$

Hence, we have

$$E\Big[\Big(\sum_{i=0}^{\eta_{l-1\wedge\theta_{y,N}}}(\xi_{l\wedge\theta_{y,N},i}-E(D_N))\Big)^2\Big|\mathcal{F}_{l-1\wedge\theta_{y,N}}\Big] \le \mathrm{Var}(D_N)\eta_{l-1\wedge\theta_{y,N}} \le c_1 Ny$$

and

$$\begin{aligned}E\Big[\Big(\sum_{i=0}^{\eta_{l-1\wedge\theta_{y,N}}}(\xi_{l\wedge\theta_{y,N},i}-E(D_N))\Big)^4|\mathcal{F}_{l-1\wedge\theta_{y,N}}\Big] &\le c_2ED_N^4\eta^2_{l-1\wedge\theta_{y,N}}+\eta_{l-1\wedge\theta_{y,N}}\\ &\le (c_2ED_N^4+1)((Ny)^2+Ny).\end{aligned}$$

Also, noting the quadratic variation estimate (3.3),

$$\frac{1}{N^2}E\left[(M(j-1\wedge\theta_{N,y})-M(k\wedge\theta_{N,y}))^2\right] \le c_3|j-k|\left(\frac{1}{N}+y\right).$$

Hence, we have, for some constant c_4 not depending on N or y, that

$$\frac{1}{N^4}E[(M([Nt]\wedge\theta_{y,N})-M([Ns]\wedge\theta_{y,N}))^4] \le c_4\max\{y^2,1\}(|t-s|^2\vee\frac{1}{N^2}).$$

Then, by Theorem 12.3 Billingsley (1999), $\{N^{-1}M([Nt]\wedge\theta_{y,N})(t)\}_{0\le t\le 1-\alpha}$ is tight for any $y<\infty$. Hence, $N^{-1}M_N(t)$ is tight in the sup topology on $[0,1-\alpha]$.

Tightness with respect to the interval $[1-\alpha,1]$, and consequently the whole interval $[0,1]$ follows similarly. □

3.3. *Asymmetric walks.* The situation is much different for asymmetric walks, in particular, the local times are of order $O(1)$, and no scaling is required. Given $p > q$, the walk starting from $x = [\alpha N]$ will exit to the right with probability tending to 1 as $N \uparrow \infty$. The sequence η_j for $1 \le j \le [(1-\alpha)N]$ is a branching process with mean offspring $ED_N = q/p < 1$ and immigration at each time of one individual. The initial population is η_1 with the distribution of $D = D_N$, a Geometric(q) random variable minus 1. Hence, this sequence is a positive recurrent Markov chain, and $\eta_{[(1-\alpha)N]}$ converges to the stationary distribution π.

On the other hand, the chain η_j for $j \ge [(1-\alpha)N]$ is the usual branching process with offspring distribution D_N (and no immigration). Hence, it dies out in finite time.

The stationary distribution π can be described by its probability generating function $\Psi(s) = \sum_{k\ge 0} \pi(k)s^k$. Let $\phi(s)$ be the probability generating function of D. Then, easy computations give that $\Psi(s) = \Psi(\phi(s))\phi(s)$.

Hence, since the distribution of $\eta_{[(1-\alpha)N]}$ converges to π, we can state a limit characterization in terms of a reversed process.

Proposition 3.5. *Consider the asymmetric walk when $p > q$ starting from $[\alpha N]$. For any $M \ge 1$, the reversed process $\{\beta_k = \eta_{[(1-\alpha)N]-k}\}_{k=0}^M$ converges in distribution to the reversed process starting from the stationary distribution π of the chain.*

However, $\{\beta_{-k} = \eta_{[(1-\alpha)N]+k}\}_{k=0}^M$ converges to a branching process with offspring distribution D starting from π.

4. Question 3: Periodicity

We now address the parity of various well-separated locations visited by the walk before exiting. We remark that different types of multiple point structures in other settings have been studied in Hamana (1997) and Pitt (1974).

Let $0 < \alpha_1 < \alpha_2 < \cdots < \alpha_k < 1$, $k \in \mathbb{Z}_+$, and $e_i \in \{0, 1\}$ for $1 \le i \le k$.

Proposition 4.1. *With respect to symmetric or weakly asymmetric walks, for $\alpha \in (0,1)$, we have*

$$\lim_{N\to\infty} P_{[\alpha N]}\left(\cap_{i=1}^k \{G([\alpha_i N]) = e_i \bmod_2\} \Big| \max_{1\le i\le k} T_{[\alpha_i N]} < \tau_N\right) = \frac{1}{2^k}.$$

In other words, in the symmetric or weakly asymmetric cases, given that the locations are visited, the parities at $\{[\alpha_i N]\}_{i=1}^k$ converge to i.i.d. fair Bernoulli random variables.

But, with respect to asymmetric walks when $p > q$, starting from $[\alpha N]$, unless $\alpha < \beta$, $[\beta N]$ is not visited with probability tending to 1. So, it makes sense only to discuss parities of sites to the right of $[\alpha N]$.

Proposition 4.2. *With respect to asymmetric walks when $p > q$, suppose $0 < \alpha < \alpha_1$. Then,*

$$\lim_{N\to\infty} P_{[\alpha N]}\left(\cap_{i=1}^k \{G([\alpha_i N]) = 1 \bmod_2\}\right) = \frac{1}{(2-(p-q))^k}.$$

By the inclusion-exclusion principle, one concludes, in the asymmetric situation, the parities at $\{[\alpha_i N]\}_{i=1}^k$ converge to i.i.d. Bernoulli random variables with success probability $(2-(p-q))^{-1}$.

We remark, with respect to the 'stochastic locker' interpretation, one concludes that the expected proportion of lockers left closed is half or $(2-(p-q))^{-1}$ times the proportion of the range in the symmetric/weakly asymmetric, or asymmetric cases respectively.

4.1. *Proofs of Propositions 4.1 and 4.2*. The proofs of the above propositions are similar. We first derive the chance a single site is left open, and then later use this development in an induction scheme. Let T_y^r be the rth hitting time of y, and $\tilde{T}_y = \inf\{n \geq 1 : X_n = y\}$ be the return time to y. The event that site y is left open, with various prescribed exits, is expressed as

$$\begin{aligned}&\{G(y) = 1 \text{ mod}_2, T_N < T_0\}\\ &\quad = \cup_{k\geq 0}\{T_y^1 < \tau_N\} \cap \{T_y^{2k+1} < \tau_N\} \cap \{T_N < T_y^{2k+2} \wedge T_0\}.\end{aligned}$$

Similarly,

$$\begin{aligned}&\{G(y) = 1 \text{ mod}_2, T_0 < T_N\}\\ &\quad = \cup_{k\geq 0}\{T_y^1 < \tau_N\} \cap \{T_y^{2k+1} < \tau_N\} \cap \{T_0 < T_y^{2k+2} \wedge T_N\}\\ &\{G(y) = 1 \text{ mod}_2\} = \cup_{k\geq 0}\{T_y^1 < \tau_N\} \cap \{T_y^{2k+1} < \tau_N\} \cap \{\tau_N < T_y^{2k+2}\}.\end{aligned}$$

Then,

$$\begin{aligned}P_x(G(y) = 1 \text{ mod}_2, T_N < T_0) &= P_x(T_y < \tau_N)P_y(T_N < \tilde{T}_y)\sum_{l\geq 0} P_y(\tilde{T}_y < \tau_N)^{2l}\\ &= \frac{P_x(T_y < \tau_N)P_y(T_N < \tilde{T}_y)}{1-(1-P_y(\tau_N < \tilde{T}_y))^2}\\ &= \frac{P_x(T_y < \tau_N)}{2-P_y(\tau_N < \tilde{T}_y)} \cdot \frac{P_y(T_N < \tilde{T}_y)}{P_y(\tau_N < \tilde{T}_y)}.\end{aligned}$$

Also,

$$\begin{aligned}P_x(G(y) = 1 \text{ mod}_2, T_0 < T_N) &= \frac{P_x(T_y < \tau_N)}{2-P_y(\tau_N < T_y)} \cdot \frac{P_y(T_0 < \tilde{T}_y)}{P_y(\tau_N < \tilde{T}_y)}\\ P_x(G(y) = 1 \text{ mod}_2) &= \frac{P_x(T_y < \tau_N)}{2-P_y(\tau_N < \tilde{T}_y)}.\end{aligned}$$

In this last expression $P_x(T_y < \tau_N)$ is the probability y is visited starting from x, and $(2-P_y(\tau_N < \tilde{T}_y))^{-1}$ is the factor specifying that y is left open. The quantity $P_y(\tau_N < \tilde{T}_y)$ can be viewed as an "escape probability."

Suppose now $x = [\alpha N]$ and $y = [\beta N]$. In the symmetric case, we compute

$$P_x(T_y < \tau_N) = \begin{cases} \frac{N-x}{N-y} & \text{for } y < x < N,\\ \frac{x}{y} & \text{for } 0 < x < y\end{cases}$$

and

$$P_y(\tilde{T}_y < \tau_N) = \frac{1}{2}P_{y-1}(T_y < T_0) + \frac{1}{2}P_{y+1}(T_y < T_N) = 1 - \frac{N}{2y(N-y)}.$$

In the (weakly) asymmetric case, we have

$$P_x(T_y < \tau_N) = \begin{cases} \frac{s_N^x - s_N^N}{s_N^y - s_N^N} & \text{for } x > y,\\ \frac{1-s_N^x}{1-s_N^y} & \text{for } x < y,\end{cases}$$

where s_N is as in (2.3) and

$$\begin{aligned} P_y(\tau_N < \tilde{T}_y) &= q_N P_{y-1}(T_0 < T_y) + p_N P_{y+1}(T_N < T_y) \\ &= \frac{p_N(1-s_N)(1-s_N^N)}{(1-s_N^y)(1-s_N^{N-y})}. \end{aligned}$$

Then,

$$P_y(\tau_N < \tilde{T}_y) \rightarrow \left\{ \begin{array}{cl} 0 & \text{for symmetric/weakly asymmetric walks} \\ p-q & \text{for asymmetric walks.} \end{array} \right.$$

Putting these observations together, we have the following result.

Proposition 4.3. *Under symmetric or weakly asymmetric motion,*

$$\begin{aligned} &\lim_{N\uparrow\infty} P_x(G(y) = 1 \text{ mod}_2 | T_y < T_N < T_0) \\ &\qquad = \lim_{N\uparrow\infty} P_x(G(y) = 1 \text{ mod}_2 | T_y < T_0 < T_N) \ = \ \frac{1}{2}, \end{aligned}$$

and hence $\lim_{N\uparrow\infty} P_x(G(y) = 1 \text{ mod}_2 | T_y < \tau_N) = 1/2$.

However, under asymmetric motion, for $x \leq y$,

$$\lim_{N\uparrow\infty} P_x(G(y) = 1 \text{ mod}_2) \ = \ \frac{1}{2-(p-q)}.$$

Proof of Proposition 4.1. Let $G_n(y) = \sum_{l=0}^{n\wedge\tau_N} 1_y(X_l)$ be the number of visits to y up to time $n \wedge \tau_N$. First, we write

$$\begin{aligned} &P_{[\alpha N]}(\cap_{i=1}^k \{T_{[\alpha_i N]} < \tau_N\}, \cap_{i=1}^k G([\alpha_i N]) = e_i \text{ mod}_2) \\ &\quad = P_{[\alpha N]}(\cap_{i=1}^k \{G([\alpha_i N]) = e_i \text{ mod}_2\}, T_{[\alpha_1 N]} < T_N < T_0) \\ &\qquad + P_{[\alpha N]}(\cap_{i=1}^k G([\alpha_i N]) = e_i \text{ mod}_2, T_{[\alpha_k N]} < T_0 < T_N). \end{aligned}$$

We now concentrate on the first term on the right when $T_N < T_0$, as the argument is similar for the second term. Since, on the set $T_N < T_0$, the walk must leave $[\alpha_1 N]$ never to return, and is also nearest-neighbor, write

$$\begin{aligned} &P_{[\alpha N]}(T_{[\alpha_1 N]} < T_N < T_0, \cap_{i=1}^k \{G([\alpha_i N]) = e_i \text{ mod}_2\}) \\ &\quad = \sum_{\substack{z_1,\ldots,z_k \\ z_1 = e_1 \text{ mod}_2}} P_{[\alpha N]}(T^{z_1}_{[\alpha_1 N]} < \tau_N, \cap_{i=2}^k \{G_{T^{z_1}_{[\alpha_1 N]}}([\alpha_i N]) = z_i\}) \\ &\qquad \cdot P_{[\alpha_1 N]}(T_{[\alpha_2 N]} < \tilde{T}_{[\alpha_1 N]} \wedge T_N) \\ &\qquad \cdot P_{[\alpha_2 N]}(T_N < T_{[\alpha_1 N]}, \cap_{i=2}^k \{G([\alpha_i N]) = e(z_i)\}) \end{aligned} \tag{4.1}$$

where $e(z_i) = e_i$ or $1 - e_i$ if z_i is even or odd respectively.

In the last factor, which deals with the parities of $k-1$ points, $[\alpha_1 N]$ can be translated to $x = 0$. Treating the limit in Proposition 4.3 as a base step, we may conclude by induction, for fixed $e(z_i)$, that

$$\lim_{N\to\infty} P_{[\alpha_2 N]}(\cap_{i=2}^k \{G([\alpha_i N]) = e(z_i)\} | T_N < T_{[\alpha_1 N]}) \ = \ 2^{-(k-1)}.$$

Hence, by bounded convergence, we may replace the last factor of (4.1) by

$$2^{-(k-1)} P_{[\alpha_2 N]}(T_N < T_{[\alpha_1 N]}) + o(1)$$

as $N \uparrow \infty$. Summing over $z_2, \dots, z_k$, we have

$$
\begin{aligned}
& P_{[\alpha N]}(T_{[\alpha_1 N]} < T_N < T_0, \cap_{i=1}^k \{G([\alpha_i N]) = e_i\}) \\
&= \left[\sum_{z_1 = e_1 \bmod_2} P_{[\alpha N]}(T^{z_1}_{[\alpha_1 N]} < \tau_N) \right] \cdot P_{[\alpha_1 N]}(T_{[\alpha_2 N]} < \tilde{T}_{[\alpha_1 N]} \wedge T_N) \\
&\qquad \cdot [2^{-(k-1)} P_{[\alpha_2 N]}(T_N < T_{[\alpha_1 N]}) + o(1)] \\
&= \frac{1}{2^{k-1}} P_{[\alpha N]}(G([\alpha_1 N]) = e_1 \bmod_2, T_{[\alpha_1 N]} < T_N < T_0) + o(1).
\end{aligned}
$$

Therefore, noting Proposition 4.3,

$$\lim_{N\to\infty} P_{[\alpha N]}(\cap_{i=1}^k \{G([\alpha_i N]) = e_i\} | T_{[\alpha_1 N]} < T_N < T_0) \ = \ \frac{1}{2^k}.$$

A similar expression is derived when the conditioning event is $T_{[\alpha_k N]} < T_0 < T_N$, and so the limit in Proposition 4.1 is recovered. □

Proof of Proposition 4.2. The proof is easier than that for Proposition 4.1. Since the probability of backtracking, $P_{[\gamma N]}(T_{[\beta N]} < \tau_N)$ is exponentially small in N for $\beta < \gamma$, and noting Proposition 4.3, we have

$$
\begin{aligned}
P_{[\alpha N]}(\cap_{i=1}^k \{G([\alpha_i N]) = 1 \bmod_2\}) &= o(1) + \prod_{i=1}^k P_{[\alpha_i N]}(G([\alpha_i N]) = 1 \bmod_2) \\
&\to (2 - (p - q))^{-k}
\end{aligned}
$$

□

References

R. F. Bass, X. Chen and J. Rosen. Moderate deviations for the range of planar random walks. *Mem. Amer. Math. Soc.* **198** (929), viii+82 (2009). MR2493313.

P. Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition (1999). ISBN 0-471-19745-9. A Wiley-Interscience Publication; MR1700749.

Y. Hamana. The fluctuation result for the multiple point range of two-dimensional recurrent random walks. *Ann. Probab.* **25** (2), 598–639 (1997). MR1434120.

G.H. den Hollander and F. Weiss. *Contemporary problems in statistical physics*, chapter Trapping in transport processes, pages 147–203. Society for Industrial and Applied Mathematics, Philadelphia (1994). ISBN 0-89871-323-4.

F. B. Knight. Random walks and a sojourn density process of Brownian motion. *Trans. Amer. Math. Soc.* **109**, 56–86 (1963). MR0154337.

J. H. Pitt. Multiple points of transient random walks. *Proc. Amer. Math. Soc.* **43**, 195–199 (1974). MR0386021.

D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition (1999). ISBN 3-540-64325-7. MR1725357.

J. Tanton. An illuminating introduction to the möbius function. *FOCUS* **27** (3), 16–17 (2007).

B. Tóth. Generalized Ray-Knight theory and limit theorems for self-interacting random walks on $\mathbf{Z}^1$. *Ann. Probab.* **24** (3), 1324–1367 (1996). MR1411497.