ON A RANDOM ENTANGLEMENT PROBLEM

GAGE BONNER, JEAN-LUC THIFFEAULT, AND BENEDEK VALKÓ

Abstract. We study a model for the entanglement of a two-dimensional reflecting Brownian motion in a bounded region divided into two halves by a wall with three or more small windows. We map the Brownian motion into a Markov Chain on the fundamental groupoid of the region. We quantify entanglement of the path with the length of the appropriate element in this groupoid. Our main results are a law of large numbers and a central limit theorem for this quantity. The constants appearing in the limit theorems are expressed in terms of a coupled system of quadratic equations.

1. Motivation

We consider a reflecting Brownian motion in a piecewise smooth bounded region of \( \mathbb{R}^2 \). This region is divided into a top and bottom half by a wall punctured by \( N \geq 3 \) small windows, as shown in Figure 1.

![Figure 1](image-url)

(A) A planar domain with \( N = 4 \) windows and a sample Brownian path. (B) The same path expressed as arcs labeled with their associated generating elements. (See Section 2.)

The Brownian path winds around the wall segments through the windows, and becomes progressively more entangled. The entanglement can be quantified by mapping the path at a given time to an element of the fundamental groupoid of the region, and considering the length of that element in the appropriate sense (see Section 2). Our goal is to study the asymptotic growth of this length as a function of time. The growth rate of words in the groupoid serves as an indication for the nature of the growth one would expect to see in the winding problem. Motivated by random walks on free groups one expects...
the length in the fundamental groupoid to grow linearly in time. In the present paper we identify formulas for both the growth rate and the limiting fluctuations around the mean, in the setting involving small windows. Our main contribution (Theorem 3.2) is the proof of these limit theorems in a general setting. The limits are described in terms of a set of coupled quadratic equations, which can be readily solved numerically.

One can follow the entanglement of the Brownian path by observing it at times when it visits a new window, and considering the length of the corresponding element in the fundamental groupoid. This fundamental groupoid is generated by equivalence classes of oriented paths connecting two windows, with each such path lying in the upper or lower half of the plane. Between successive observations, the groupoid element corresponding to the path is appended by a random generating element whose distribution depends on the location inside the window. Motivated by the narrow escape problem [13], as the windows shrink in size this location dependence disappears, and we arrive at a Markov chain on the fundamental groupoid. Our limit theorems are about the length of the groupoid element in this Markov chain.

Probabilistic winding problems on surfaces have a long history. A classical example is the asymptotic behavior of the winding of a planar Brownian motion around a point. Spitzer [32] showed that the winding angle at time $t$, scaled by $\log t$, converges to a standard Cauchy distribution as time goes to infinity. The fact that the limit distribution has no moments can be explained by the large amount of winding that the Brownian path can pick up when it comes near the origin. This model has been thoroughly investigated by many authors [22, 24–26].

When using Brownian motion to model, say, polymer entanglement [10], it is more realistic to regularize the problem in some way. This can be accomplished, for example, by replacing the punctual winding center by a finite topological disk [8, 10], by adding a persistence length to the motion [33], or by considering a random walk instead [1–4, 27, 28]. In the regularized problem the scaling limit for the winding angle becomes the hyperbolic secant distribution, where all the moments exist. Unsurprisingly, confinement to a finite region greatly increases the rate of winding, since the Brownian path returns near the winding center more frequently [8, 10, 35].

A more challenging problem is the study of winding around multiple points or topological disks. A natural first approach to this problem is the homological route, where one examines the joint distribution of winding angles around each winding center. In the scaling limit these winding angles converge to independent Cauchy distributions [24, 25]. The homological route is inherently Abelian, in that the order of winding around the centers is lost. Watanabe [34] studied winding on punctured surfaces of higher genus, and derived Gaussian limit distributions for the windings around each handle.

Another approach is via the fundamental group of the punctured surface, which is the group of deck transformations on its universal cover. In that case, the non-Abelian aspect
of the windings is captured, and we may regard distance in the universal cover as a measure of entanglement of the Brownian motion. This approach was first introduced by Itô and McKean [14,20] who considered the twice-punctured plane. (See also [19,21].) Gruet [11] finds that the length of the word at time $t$ in the fundamental group of the thrice-punctured sphere grows at least like $t \log t$ as $t \to \infty$. Desenonges [6] considers a similar problem on a wider class of surfaces with $n$ punctures. See also the book by Nechaev [23] for winding in an infinite lattice of points. Note that the region in our Brownian entanglement problem is topologically equivalent to a sphere with $N$ holes, hence our result belongs to this class of non-Abelian problems. Our Markov chain can also be considered as a random walk on a regular language (see Remark 3.3).

Our paper is organized as follows. Section 2 and Section 3 contain the precise setup of the problem and our main result. In Section 4 we give a number of applications of our main theorem. Section 5 provides the key steps of the proof, which are proved in the rest of the paper (Section 6 and Appendix A).

2. Preliminaries

2.1. The fundamental groupoid $G_N$. We consider the groupoid representing the homotopy classes of continuous paths that start and end at the midpoints of the windows as in Figure 1.

Recall that a groupoid is defined by a set of ‘objects’ $O$ and ‘arrows’ $X$ and the following functions:

(G1) There are functions $s$ (source) and $t$ (target) from $X \to O$.

(G2) There is a composition function $(f_1,f_2) \to f_1f_2$ on a subset of $X \times X$ which is defined for $f_1,f_2$ if $t(f_1) = s(f_2)$, and in that case $s(f_1f_2) = s(f_1)$, $t(f_1f_2) = t(f_2)$. The composition function is associative.

(G3) For each $i \in O$ there is a unique unit element $e_i \in X$ with $s(e_i) = t(e_i) = i$ for which $e_if = f$, $fe_i = f$ whenever these are defined.

(G4) There is an inverse for each element of $X$ satisfying $s(f^{-1}) = t(f)$, $t(f^{-1}) = s(f)$ and $ff^{-1} = e_{s(f)}$, $f^{-1}f = e_{t(f)}$.

We consider a groupoid $G_N$ with object set $O_N = \{1,2,\ldots,N\}$, and arrow set $X_N = X$ generated by the elements in $A_N := \{A_{i,j}^{(k)} : i \neq j, \ 1 \leq i,j \leq N, \ k \in \{-1,1\}\}$,

\begin{equation}
A_N := \{A_{i,j}^{(k)} : i \neq j, \ 1 \leq i,j \leq N, \ k \in \{-1,1\}\},
\end{equation}

with

\begin{equation}
\begin{align*}
s(A_{i,j}^{(k)}) &= i, \\
t(A_{i,j}^{(k)}) &= j.
\end{align*}
\end{equation}

For convenience we define $A_{i,i}^{(k)} := e_i$. The set of generating relations for our groupoid is given by the relations

\begin{equation}
A_{i,j}^{(k)}A_{j,\ell}^{(k)} = A_{i,\ell}^{(k)} \quad i,j,\ell \in O, \ k \in \{-1,1\}.
\end{equation}

We call a composition of finitely many generating elements a word. We include the unit elements as words, and call them empty words. Each arrow in $G_N$ corresponds to an equivalence class of words. Two words are in the same equivalence class if they can be transformed into each other by repeated application of relations $(2.3)$. We say that a nonempty word is reduced if we cannot apply the relations $(2.3)$ to reduce the number of generators used in the word. We consider the empty words reduced as well.

For $i \neq j$, the arrow $A_{i,j}^{(+1)}$ (respectively, $A_{i,j}^{(-1)}$) corresponds to a simple path in the upper (respectively, lower) part of the region connecting window $i$ to window $j$ as in Figure 1b. Figure 2 shows a schematic of the groupoid structure for $N = 3$.

The arrows of the groupoid $G_N$ can be represented as equivalence classes of paths in the directed multi-graph which has vertices $\mathbb{O} = \{1, 2, \ldots, N\}$ and directed edges of the form $(i, j, k)$ with $i \neq j \in \mathbb{O}$, $k \in \{-1, 1\}$. The generating elements $A_{i,j}^{(k)}$ correspond to the directed edges $(i, j, k)$; a composition of generating elements (a word) corresponds to a path in the multi-graph. The starting and ending vertices of a path are the results of the source and target functions. Two paths are in the same equivalence class (and correspond to the same arrow in $X$) if they can be transformed into each other by the repeated use of the following operations and their inverses

(EC1) Deleting a backtracking step $(i, j, k)$, $(j, i, k)$
(EC2) Replacing two consecutive steps $(i, j, k)$, $(j, \ell, k)$ with $(i, \ell, k)$ if $i, j, \ell$ are different.

These operations correspond to the generating relations $(2.3)$. A path corresponds to a reduced word if if we cannot use either of the moves [EC1] and [EC2] on it. An important consequence of the properties of our groupoid is that each non-unit arrow can be uniquely represented as a product of elements of $A_N$ which alternate between $\pm 1$ in the upper index $(k)$. More precisely, we have the following lemma which is proved in Appendix A.

**Lemma 2.1.** Each arrow $w \in X$ can be represented as a reduced word in a unique way. This reduced word is either an empty word or, for some $d \geq 1$, a product of the form

$$w = A_{i_1, i_2}^{(k_1)} A_{i_2, i_3}^{(-k)} A_{i_3, i_4}^{(k)} \cdots A_{i_{d}, i_{d+1}}^{((-1)^{d+1}k)}$$

such that $i_\ell \neq i_{\ell+1}$ for all $1 \leq \ell \leq d$. (2.4)

We say that $|\cdot|_L : X \to [0, \infty)$ is a metric on $X$ generated by $A_N$ if $|e_i|_L = 0$ for all $i$ and for any nonempty $w \in X$ we have

$$|w|_L = \sum_{\ell=1}^{d} \left| A_{i_\ell, i_{\ell+1}}^{(k_\ell)} \right|_L,$$

where $\prod_{\ell=1}^{d} A_{i_\ell, i_{\ell+1}}^{(k_\ell)}$ is the unique reduced representation of $w$ given by Lemma 2.1. We reserve $|\cdot|$ to denote the number of generators in the reduced representation of $w \in X$;

$$|w| = \prod_{\ell=1}^{d} \left| A_{i_\ell, i_{\ell+1}}^{(k_\ell)} \right| = d.$$ (2.6)
2.2. The Markov chain on $G_N$. We study discrete time Markov chains $\{W_n\}_{n \geq 0}$ on the arrow set $X$. We assume the following:

**Assumption 2.2.** In each step the value of $W_n$ changes by the right composition of a generating element $A_{i,j}^{(k)}$, i.e. $W_n^{-1}W_{n+1} \in A_N$ for all $n \geq 0$.

**Assumption 2.3.** The conditional probability of the increments given the starting state depends on the starting state only through its target.

These two assumptions imply that for $x, y \in X$ the transition probability function is of the form

$$P(W_{n+1} = y \mid W_n = x) := \begin{cases} p_{i,j}^{(k)} & \text{if } x^{-1}y = A_{i,j}^{(k)} \text{ and } i \neq j \\ 0 & \text{otherwise} \end{cases} \quad (2.7)$$

We further assume:

**Assumption 2.4.** The jump probabilities satisfy $p_{i,j}^{(k)} \in (0, 1)$ for each $(i, j, k)$.

We consider $N \geq 3$ since, if $N = 2$, then the process $\{W_{2n}\}_{n \geq 0}$ reduces to the Abelian case, namely a Markov chain on the free group of rank 1, namely a lazy random walk on $\mathbb{Z}$.

From Eq. (2.7), we see that there are $2N(N-1)$ transition probabilities and we must have $\sum_{j,k} p_{i,j}^{(k)} = 1$ for each $1 \leq i \leq N$. We call $W_n$ the “word” at time $n$. We use the notation $P_x(\cdot)$ and $E_x[\cdot]$ to indicate that the probabilities and expectations in question are calculated under the initial condition $W_0 = x$. 

---

**Figure 2.** The arrows represented by members $A_{i,j}^{(k)}$ of $A_3$ with $i < j$. The objects in our groupoid $\{1, 2, 3\}$ are represented by circles. Each arrow points from $s(A_{i,j}^{(k)})$ to $t(A_{i,j}^{(k)})$. 

Define the sets
\[ X_i := \{ w : s(w) = i, w \in X \}, \quad 1 \leq i \leq N. \] (2.8)

Some important consequences of our assumptions are collected in the following lemma which is proved in Appendix A.

**Lemma 2.5.**

(i) For any \( x \in X \), under \( P_x \) the process \( \{x^{-1}W_n\}_{n \geq 0} \) has the same distribution as the process \( \{W_n\}_{n \geq 0} \) under \( P_{e(x)} \).

(ii) There is a positive probability path in the Markov chain \( \{W_n\}_{n \geq 0} \) between two words \( w_1 \neq w_2 \) if and only if \( w_1, w_2 \in X_i \) for some \( 1 \leq i \leq N \).

(iii) Suppose that \( \{v_n\}_{1 \leq n \leq n'} \) is such a path where \( v_1 = w_1 \) and \( v_{n'} = w_2 \). Consider the reduced composition
\[ w_1^{-1}w_2 = \prod_{\ell=1}^{d} A_{i,\ell,j,\ell}^{(k_{\ell})} \quad d \geq 1 \] (2.9)
given by Lemma 2.1. Then, there exist times \( \{n_m\}_{1 \leq m \leq d-1} \) satisfying \( 1 < n_1 < n_2 < \cdots < n_{d-1} < n' \), such that \( v_{n_m} = w_1 \prod_{\ell=1}^{m} A_{i,\ell,j,\ell}^{(k_{\ell})} \) for each \( 1 \leq m \leq d-1 \).

![Figure 3](image-url)

**Figure 3.** An illustration for Statement (iii) of Lemma 2.5. Let \( w_1 = A_{1,2}^{(1)} \) (red, dashed) and \( w_1^{-1}w_2 = A_{2,5}^{(-1)} A_{5,3}^{(1)} A_{3,4}^{(-1)} \) (blue, solid) in the \( N = 5 \) case. A path joining \( w_1 \) and \( w_2 \) must visit the intermediate words \( w_1 A_{2,5}^{(-1)}, w_1 A_{5,3}^{(-1)} A_{3,4}^{(1)} \) in this order.

**Remark 2.6.** To connect with the original problem described in Section 1, we first note that any path in the domain \( D \) (described in Figure 1a) starting at a window and ending at a window can be naturally mapped into an element of \( X \). Indeed, an element of \( X \) corresponds to the homotopy class of a path connecting two windows, where we allow the starting and ending points of the path to move inside a window.

Let \( B_t \) be the position of the reflected Brownian motion in \( D \) at time \( t \). Let \( \tau_0 = \inf_{s \geq 0} \{ B_s \text{ at a window} \} \) and
\[ \tau_{n+1} = \inf_{s \geq \tau_n} \{ B_s \text{ at a window different from the window visited at time } \tau_n \}. \] (2.10)
Denoting by $W_n$ the element in $X$ corresponding to the Brownian path in the time interval $[\tau_0, \tau_n]$, we see that the sequence $(W_n, B_{\tau_n})_{n \geq 0}$ is a Markov chain on $X \times D$. By the narrow escape problem [13] the transition distribution of this Markov chain will only depend on the first coordinate (an element in $X$) as all window sizes go to 0, which leads to a Markov chain on $X$ satisfying the conditions described in Section 2.2. The narrow escape problem also implies that understanding the growth of the length of $W_n$ in $n$ allows us to understand the growth of the entanglement of the Brownian path in $t$.

3. Statement of main result

We now state our main result, which is the computation of the almost-sure limit $\lim_{n \to \infty} |W_n|_L/n$ as well as a central limit theorem. To do this, we require a set of functions whose properties are collected in the following proposition.

**Proposition 3.1.** There is an $\epsilon > 0$ and a unique set of complex functions

$$\{R_{i,j}^{(k)} : D_{1+\epsilon} \to \mathbb{C} : i \neq j, \ 1 \leq i, j \leq N, \ k \in \{-1, 1\}\}, \quad D_r = \{z : |z| < r\}, \quad (3.1)$$

satisfying the following properties:

(R1) Each $R_{i,j}^{(k)}$ is complex analytic in the disk $D_{1+\epsilon}$.

(R2) If $\lambda \in [0, 1]$, then $R_{i,j}^{(k)}(\lambda) \in (0, 1)$.

(R3) They satisfy the system of equations

$$R_{i,j}^{(k)} = \lambda \left[ p_{i,j}^{(k)} + \sum_{m \neq i,j} p_{i,m}^{(k)} R_{m,j}^{(k)} + \sum_{m \neq i} p_{i,m}^{(-k)} R_{m,i}^{(-k)} R_{i,j}^{(k)} \right], \quad \text{for } \lambda \in D_{1+\epsilon}. \quad (3.2)$$

Let $B : \{-1, 1\} \times D_{1+\epsilon} \times \mathbb{C} \to \mathbb{C}^{N \times N}$ be a matrix-valued function whose $i,j$th entry is

$$B_{i,j}(k; \lambda, z) := (1 - \delta_{i,j}) z |A_{i,j}^{(k)}| \cdot R_{i,j}^{(k)}(\lambda) \quad (3.3)$$

where $\delta_{i,j}$ is the Kronecker delta. Now let

$$h(\lambda, z) := \det[\mathbf{I} - B(1; \lambda, z) B(-1; \lambda, z)], \quad (3.4)$$

where $\mathbf{I}$ is the identity matrix. Let

$$\gamma := \frac{\partial \lambda h(1, 1)}{\partial \lambda h(1, 1)}; \quad (3.5a)$$

$$\sigma^2 := \frac{\partial^2 h(1, 1) + \partial \lambda h(1, 1) - 2\gamma \partial^2 h(1, 1) + \gamma^2 (\partial^2 h(1, 1) + \partial \lambda h(1, 1))}{\partial \lambda h(1, 1)}. \quad (3.5b)$$

We are now ready to state our main result:

**Theorem 3.2.** Consider the Markov chain $\{W_n\}_{n \geq 0}$ satisfying Assumptions 2.2, 2.3 and 2.4 and having transition probabilities defined in Eq. (2.7). Then, for any initial condition
We have
\[
\lim_{n \to \infty} \frac{|W_n|_L}{n} = \gamma \quad \text{a.s.}
\]  
(3.6a)
\[
\frac{|W_n|_L - \gamma n}{\sqrt{n}} \to N(0, \sigma^2) \quad \text{in law.}
\]  
(3.6b)

The constants \(\gamma\) and \(\sigma^2\) are defined in Eq. (3.5) and \(N(0, \sigma^2)\) denotes the normal distribution with mean 0 and variance \(\sigma^2\).

**Remark 3.3.** Gilch [9] studies similar problems in the context of random walks on regular languages. A random walk on a regular language as defined in [17] is a Markov chain on the set of all finite words from a finite alphabet with the following conditions. In one jump only the last two letters of a word may be modified and at most one letter may be adjoined or deleted. The transition probabilities only depend on the last two letters of the current word. Our process \(\{W_n\}_{n \geq 0}\) is a random walk on a regular language formed by the alphabet \(A_N\).

Theorem 2.4 in [9] provides a law of large numbers under the assumption that the considered random walk is transient. This could potentially lead to another way to obtain the constant \(\gamma\) in Theorem 3.2. However, the identification of the constant appearing in Theorem 2.4 of [9] requires the solution of a more complicated problem than in our case.

4. **Examples**

Here we demonstrate several applications of Theorem 3.2. We will consider two simple metrics and compute the constants in Eqs. (3.5a) and (3.5b). The first is the metric \(|\cdot|\) defined in Eq. (2.6); the second is the metric \(|\cdot|_F\) generated by \(A_N\) defined for the generators as
\[
|A^{(k)}_{i,j}|_F = |i - j|.
\]  
(4.1)

4.1. **The one parameter case with** \(N = 3\). We take \(N = 3\) and the set of transition probabilities
\[
p^{(1)}_{2,1} = p^{(1)}_{2,3} = p^{(-1)}_{2,1} = p^{(-1)}_{2,3} = 1/4, \\
p^{(1)}_{1,2} = p^{(1)}_{3,2} = p^{(-1)}_{1,2} = p^{(-1)}_{3,2} = q, \\
p^{(1)}_{1,3} = p^{(1)}_{3,1} = p^{(-1)}_{1,3} = p^{(-1)}_{3,1} = 1/2 - q,
\]
for \(0 < q < 1/2\). This represents a situation where the planar domain of Section 1 is left-right and up-down symmetric about its center. Let \(\gamma_3(q), \sigma^2_3(q)\) be the constants appearing in Theorem 3.2 for the \(|\cdot|\) metric and let \(\gamma_{3,F}(q), \sigma^2_{3,F}(q)\) be the same constants for the \(|\cdot|_F\)
We define
\[ \sigma_Q = \frac{3q + (1 - 4q)\sqrt{(8 - 7q)q}}{4(1 - 4q^2)}, \] (4.2a)
\[ \gamma_{3, f}(q) = \frac{q}{2(2q + 1)}, \] (4.2b)
\[ \sigma_3^2 = \frac{4(8+5Q)+(68-56Q)q+(500-101Q)q^2-(1471+64Q)q^3+8(1+42Q)q^4+728q^5}{8(1+2q)^4(8-23q+14q^2)}, \] (4.2c)
\[ \sigma_3^2, f(q) = \frac{32+4Q+(36+28Q)q+8(80-21Q)q^2-(199+30Q)q^3+70q^4}{2(1+2q)^4Q^2}, \] (4.2d)

where \( Q = \sqrt{(8 - 7q)q} \). The constants \( \gamma_3(q), \gamma_{3, f}(q) \) are plotted in Figure 4 and the constants \( \sigma_3^2(q), \sigma_3^2, f(q) \) are plotted in Figure 5.

By our choice of probabilities, Eq. (3.2) reduces to a set of three equations by symmetry. We define
\[ R_1 := R_{2,1}^{(1)} = R_{2,3}^{(1)} = R_{2,1}^{(-1)} = R_{2,3}^{(-1)}, \]
\[ R_2 := R_{1,2}^{(1)} = R_{3,2}^{(1)} = R_{1,2}^{(-1)} = R_{3,2}^{(-1)}, \]
\[ R_3 := R_{1,3}^{(1)} = R_{3,1}^{(1)} = R_{1,3}^{(-1)} = R_{3,1}^{(-1)}, \]

which leads to
\[ R_1 = \frac{1}{4}\lambda[1 + R_3 + 2R_2R_1], \] (4.3a)
\[ R_2 = \lambda \left[q + \left(\frac{1}{2} - q\right)R_2 + qR_1R_2 + \left(\frac{1}{2} - q\right)R_3R_2\right], \] (4.3b)
\[ R_3 = \lambda \left[\frac{1}{2} - q + qR_1 + qR_1R_3 + \left(\frac{1}{2} - q\right)R_3\right]. \] (4.3c)

Substituting \( \lambda = 1 \) gives a cubic equation for \( R_2(1) \). We choose the solution which satisfies \( R_2(1) \in (0, 1) \) (by property (R2)). By implicit differentiation of Eq. (4.3) it follows that
\[ R_1(1) = \frac{1}{2}; \quad R'_1(1) = \frac{3q + 2 + Q}{2(Q - q)}, \] (4.4a)
\[ R_2(1) = \frac{3q - Q}{2(2q - 1)}; \quad R'_2(1) = \frac{2(q + 1)}{Q}, \] (4.4b)
\[ R_3(1) = \frac{q - 2 + Q}{2(2q - 1)}; \quad R'_3(1) = \frac{2Q + q(Q - 6 + q(5 - 4q - 4Q))}{q(2q - 1)(Q + 7q - 8)}. \] (4.4c)
\[ R''_1(1) = \frac{4Q+16(2+Q)q+2(42+19Q)q^2(26+31Q)q^3-3(53+4Q)q^4+84q^5}{4(1-q)^3q^2Q^2} \] (4.4d)
\[ R''_2(1) = \frac{8(3+Q)+(25Q-12)q+(49+4Q)q^2-4(9+7Q)q^3-16q^4}{(1-q)Q}, \] (4.4e)
\[ R''_3(1) = \frac{-16(1+Q)+(48-34Q)q+6(23Q-43)q^2+2(329+18Q)q^3-(176+137Q)q^4+(28Q-541)q^5+300q^6}{2(1-q)^3(2q-1)Q^3} \] (4.4f)
Since our $p_{i,j}^{(k)}$ are symmetric with respect to $k = \pm 1$, in the $|\cdot|$ metric, we construct

$$
B(1; \lambda, z) = B(-1; \lambda, z) = \begin{pmatrix} 0 & zR_2 & zR_3 \\ zR_1 & 0 & zR_1 \\ zR_2 & zR_3 & 0 \end{pmatrix},
$$

so that

$$
h(\lambda, z) = (1 - z^2 R_3^2)[(1 - 2z^2 R_1 R_2)^2 - z^2 R_3^2].
$$

Substituting this into Eqs. (3.5a) and (3.5b) and using Eq. (4.4) gives the required constants. Similarly, in the $|\cdot|_F$ metric, we have

$$
B(1; \lambda, z) = B(-1; \lambda, z) = \begin{pmatrix} 0 & zR_2 & z^2 R_3 \\ zR_1 & 0 & zR_1 \\ z^2 R_2 & zR_3 & 0 \end{pmatrix}.
$$

Hence, we find

$$
h(\lambda, z) = (1 - z^4 R_3^2)[(1 - 2z^2 R_1 R_2)^2 - z^4 R_3^2].
$$

We again substitute this into Eqs. (3.5a) and (3.5b) and use Eq. (4.4) to obtain the required constants.

![Figure 4](image-url) **Figure 4.** The constants $\gamma_3(q)$ of Eq. (4.2a) and $\gamma_{3,F}(q)$ of Eq. (4.2b). We note that $\gamma_3(q)$ has a maximum value of $1/4$ when $q = 1/4$ and $\gamma_{3,F}(q)$ has a maximum value of $(2/23)(2\sqrt{6} - 1)$ when $q = (8 - \sqrt{6})/29$. 
4.2. An asymmetric case with \( N = 3 \). We take \( N = 3 \) and the following arbitrarily chosen set of probabilities:

\[
\begin{align*}
p_{2,1}^{(1)} &= 17/40; \quad p_{2,3}^{(1)} = 1/5; \quad p_{2,1}^{(-1)} = 1/8; \quad p_{2,3}^{(-1)} = 1/4 \\
p_{1,2}^{(1)} &= 43/70; \quad p_{3,2}^{(1)} = 43/72; \quad p_{1,2}^{(-1)} = 1/7; \quad p_{3,2}^{(-1)} = 1/8 \\
p_{1,3}^{(1)} &= 1/10; \quad p_{3,1}^{(1)} = 1/9; \quad p_{1,3}^{(-1)} = 1/7; \quad p_{3,1}^{(-1)} = 1/6.
\end{align*}
\]

Truncating to six significant digits, solving Eq. (3.2) numerically gives, for \( r_{i,j}^{(k)} := R_{i,j}^{(k)}(1) \),

\[
\begin{align*}
r_{2,1}^{(1)} &= 0.591572; \quad r_{2,3}^{(1)} = 0.404666; \quad r_{2,1}^{(-1)} = 0.388890; \quad r_{2,3}^{(-1)} = 0.579542 \\
r_{1,2}^{(1)} &= 0.769190; \quad r_{3,2}^{(1)} = 0.791039; \quad r_{1,2}^{(-1)} = 0.305398; \quad r_{3,2}^{(-1)} = 0.245890 \\
r_{1,3}^{(1)} &= 0.386687; \quad r_{3,1}^{(1)} = 0.538119; \quad r_{1,3}^{(-1)} = 0.387184; \quad r_{3,1}^{(-1)} = 0.300936,
\end{align*}
\]

and, for \( d_{i,j}^{(k)} := dR_{i,j}^{(k)}(\lambda)/d\lambda|_{\lambda=1} \),

\[
\begin{align*}
d_{2,1}^{(1)} &= 1.36978; \quad d_{2,3}^{(1)} = 1.37284; \quad d_{2,1}^{(-1)} = 2.05937; \quad d_{2,3}^{(-1)} = 2.69097 \\
d_{1,2}^{(1)} &= 1.44102; \quad d_{3,2}^{(1)} = 1.71828; \quad d_{1,2}^{(-1)} = 1.31219; \quad d_{3,2}^{(-1)} = 0.991008 \\
d_{1,3}^{(1)} &= 1.56411; \quad d_{3,1}^{(1)} = 1.99059; \quad d_{1,3}^{(-1)} = 2.01524; \quad d_{3,1}^{(-1)} = 1.19855,
\end{align*}
\]

and, for \( v_{i,j}^{(k)} := d^2 R_{i,j}^{(k)}(\lambda)/d\lambda^2|_{\lambda=1} \),

\[
\begin{align*}
v_{2,1}^{(1)} &= 10.3365; \quad v_{2,3}^{(1)} = 12.8916; \quad v_{2,1}^{(-1)} = 26.2278; \quad v_{2,3}^{(-1)} = 32.1490 \\
v_{1,2}^{(1)} &= 9.45100; \quad v_{3,2}^{(1)} = 13.7182; \quad v_{1,2}^{(-1)} = 15.3088; \quad v_{3,2}^{(-1)} = 11.6415 \\
v_{1,3}^{(1)} &= 15.5010; \quad v_{3,1}^{(1)} = 18.8416; \quad v_{1,3}^{(-1)} = 26.1337; \quad v_{3,1}^{(-1)} = 14.3857.
\end{align*}
\]
Here, $\mathbf{B}(1; \lambda, z)$ and $\mathbf{B}(-1; \lambda, z)$ will be different; however they are each $3 \times 3$ matrices with the same structure as Eqs. (4.5) and (4.7) in the corresponding metric. Via Eqs. (3.5a) and (3.5b) we find the constants $\gamma_a, \sigma_a^2$ in the $\cdot |$ metric and $\gamma_{a,F}, \sigma_{a,F}^2$ in the $\cdot |_F$ metric to be

$$\begin{align*}
\gamma_a &= 0.272913 \ldots \quad \sigma_a^2 = 0.587598 \ldots \\
\gamma_{a,F} &= 0.334211 \ldots \quad \sigma_{a,F}^2 = 0.916276 \ldots
\end{align*}$$

### 4.3. The totally symmetric case.

For any $N \geq 3$, we take $p_{i,j}^{(k)} = 1/(2N - 2)$ for all $(i,j,k)$. Let $\gamma_{\text{sym}}, \sigma_{\text{sym}}^2$ be the constants appearing in Theorem 3.2 for the $\cdot |$ metric and let $\gamma_{\text{sym},F}, \sigma_{\text{sym},F}^2$ be the same constants for the $\cdot |_F$ metric. We will show that

$$\begin{align*}
\gamma_{\text{sym}} &= \frac{N - 2}{2(N - 1)}, \\
\gamma_{\text{sym},F} &= \frac{(N + 1)(N - 2)}{6(N - 1)}, \\
\sigma_{\text{sym}}^2 &= \frac{N^2 + 2N - 4}{4(N - 1)^2}, \\
\sigma_{\text{sym},F}^2 &= \frac{11N^5 - 2N^4 + 15N^3 - 36N - 8}{180N(N - 1)^2}.
\end{align*}$$

Note that Eqs. (4.10a) and (4.10b) also hold for $N = 2$ since $\{W_n\}_{n \geq 0}$ is recurrent in this case.

We will use the following lemma for the characteristic polynomial of a Kac–Murdock–Szegő matrix [15].

**Lemma 4.1.** Let $\mathbf{U}_n(z)$ be an $n \times n$ matrix whose $i,j$th entry is $[\mathbf{U}_n(z)]_{i,j} = z^{|i-j|}$. Let $\phi_n(x, z) = \det[\mathbf{U}_n(z) - x\mathbf{I}]$. Then, defining $\phi_0(x, z) = 1$, we have

$$\phi_n(x, z) = (1-x-z^2(1+x))\phi_{n-1}(x, z) - x^2z^2\phi_{n-2}(x, z); \phi_0(x, z) = 1, \phi_1(x, z) = 1-x. \quad (4.11)$$

**Proof of Lemma 4.1.** We multiply the second row of $\mathbf{U}_n(z) - x\mathbf{I}$ by $-z$ and add it to the first row. Then, we multiply the second column of the resulting matrix by $-z$ and add it to the first column. The result is that

$$\phi_n(x, z) = \begin{vmatrix}
1 - x & z & z^2 & \ldots \\
z & 1 - x & z & \ldots \\
z^2 & z & 1 - x & \ldots \\
\vdots & \vdots & \vdots & \ddots
\end{vmatrix} = \begin{vmatrix}
1 - x - z^2(1+x) & xz & 0 & \ldots \\
xz & 1 - x & z & \ldots \\
0 & z & 1 - x & \ldots \\
\vdots & \vdots & \vdots & \ddots
\end{vmatrix}. \quad (4.12)$$
Expanding the determinant along the first column gives

\[
\phi_n(x, z) = (1 - x - z^2(1 + x))\phi_{n-1}(x, z) - xz \begin{vmatrix} xz & 0 & 0 & \ldots \\ z & 1 - x & z & \ldots \\ z^2 & z & 1 - x & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{vmatrix} \phi_{n-1}(x, z) - xz \begin{vmatrix} 1 - x & z & \ldots \\ z & 1 - x & \ldots \\ \vdots & \ddots & \ddots \end{vmatrix} \phi_{n-2}(x, z).
\]

(4.13)

Since the transition probabilities are all the same for a given \(N\), Eqs. (3.2) are invariant under interchange of any pair \((i_1, j_1, k_1), (i_2, j_2, k_2)\). Therefore \(R_{\text{sym}}(\lambda) := R_{i,j}(\lambda)\) satisfies a quadratic equation,

\[
R_{\text{sym}} = \frac{\lambda}{2(N-1)} \left[ 1 + (N-2)R_{\text{sym}} + (N-1)R_{i,j}^2 \right].
\]

(4.14)

Substituting \(\lambda = 1\) and taking the root such that \(R_{\text{sym}}(1) \in (0, 1)\) gives

\[
R_{\text{sym}}(1) = \frac{1}{N-1}, \quad R'_{\text{sym}}(1) = \frac{2}{N-2}, \quad R''_{\text{sym}}(1) = \frac{4(N^2 - 2)}{(N-2)^3}.
\]

(4.15)

In the \(|·|\) metric, \(B_{i,j}(k; \lambda, z) = (1 - \delta_{i,j})zR_{\text{sym}}(\lambda)\). Since \(B(k; \lambda, z)\), does not depend on \(k\), we will write \(H(\lambda, z) := B(\lambda, z)\). By Lemma 4.1,

\[
det[I \pm H(\lambda, z)] = (\pm zR_{\text{sym}}(\lambda))^N \phi_N \left( 1 \mp \frac{1}{zR_{\text{sym}}(\lambda)}, 1 \right).
\]

(4.16)

Hence, we compute \(h(\lambda, z) = det[I + H(\lambda, z)]det[I - H(\lambda, z)]\). This can be computed up to terms of order \((z - 1)^2\) by substituting \(z = 1\) into Eq. (4.11), solving the resulting linear recurrence relation and applying implicit differentiation. Via Eqs. (3.5a) and (3.5b) we find the constants \(\gamma_{\text{sym}}, \sigma_{\text{sym}}^2\) by expanding \(\phi_N(x, z)\) in powers of \((z - 1)\).

In the \(|·|_F\) metric, \(B_{i,j} = (1 - \delta_{i,j})z^{i+j}R_{\text{sym}}\). By Lemma 4.1,

\[
det[I \pm H(\lambda, z)] = (\pm R_{\text{sym}}(\lambda))^N \phi_N \left( 1 \mp \frac{1}{R_{\text{sym}}(\lambda)} , z \right).
\]

(4.17)

We again have \(h(\lambda, z) = det[I + H(\lambda, z)]det[I - H(\lambda, z)]\) and the same arguments as above give \(\gamma_{\text{sym},F}, \sigma_{\text{sym},F}^2\).

Remark 4.2. In the \(|·|\) metric, due to the symmetry of the problem, the process \(\{|W_n|\}_{n \geq 0}\) is a lazy nearest neighbor walk on the non-negative integers with transition probabilities

\[
P(|W_{n+1} = y | |W_n| = x) = \begin{cases} 1 & \text{if } x = 0 \text{ and } y = 1 \\ \frac{N-1}{2(N-1)} & \text{if } x \neq 0 \text{ and } y = x + 1 \\ \frac{N+1}{2(N+1)} & \text{if } x \neq 0 \text{ and } y = x \\ 0 & \text{if } x \neq 0 \text{ and } y = x - 1 \end{cases}
\]

(4.18)
This process is transient so Eqs. (4.10a) and (4.10c) follow by direct computation.

5. Outline of proof

Our proof strategy uses the double generating function method of Sawyer and Steger (29, Theorem 2.2):

**Theorem 5.1** (Sawyer and Steger). Let \( \{Y_n\}_{n \geq 0} \) be a sequence of non-negative random variables and suppose that we can write for some \( \delta > 0 \)

\[
G(\lambda, z) := E\left[ \sum_{n=0}^{\infty} z^n Y_n \lambda^n \right] = \frac{C(\lambda, z)}{g(\lambda, z)} \quad \text{for } \lambda, z \in (1 - \delta, 1),
\]

(5.1)

where \( C(\lambda, z), g(\lambda, z) \) can be extended as analytic functions to the regions \( 1 - \delta < |\lambda| < 1 + \delta, |z - 1| < \delta \) in the complex plane, and \( C(1, 1) \neq 0 \). Let

\[
\mu := \frac{\partial_z g(1, 1)}{\partial_{\lambda} g(1, 1)},
\]

(5.2a)

\[
\nu^2 := \frac{\partial^2_z g(1, 1) + \partial_z g(1, 1) - 2\mu \partial^2_{\lambda z} g(1, 1) + \mu^2 (\partial^2_{\lambda} g(1, 1) + \partial_{\lambda} g(1, 1))}{\partial_{\lambda} g(1, 1)}.
\]

(5.2b)

Then

\[
\lim_{n \to \infty} \frac{Y_n}{n} = \mu \quad \text{a.s.} \quad (5.3a)
\]

\[
\frac{Y_n - \mu n}{\sqrt{n}} \to \mathcal{N}(0, \nu^2) \quad \text{in law.} \quad (5.3b)
\]

We will apply Theorem 5.1 with \( Y_n = |W_n|_L \). In order to understand the expectation in Eq. (5.1) we introduce a family of stopping times. For \( x \in \mathbb{X} \) let

\[
T(m, x) := \inf_{k \geq 0} \{W_{m+k} = W_m x\};
\]

(5.4)

note that these can be \( \infty \). We define a set of \( 2N(N - 1) \) generating functions, one for each element of \( \mathbb{A}_N \) as

\[
R_{ij}^{(k)}(\lambda) := E_{\omega} \left[ \lambda^{T(0,A^{(k)})_{ij}} \right],
\]

(5.5)

defined for those \( \lambda \in \mathbb{C} \) where the expectation is finite. As we will show, these generating functions are the functions described in Proposition 3.1. To do this, we will show that the \( R_{ij}^{(k)}(\lambda) \) introduced in Eq. (5.5) uniquely satisfy Properties (R1), (R2) and (R3) of Proposition 3.1.

**Remark 5.2.** Since each \( R_{ij}^{(k)}(\lambda) \) is a power series whose coefficients are non-negative numbers summing to at most 1, we immediately see that they satisfy weaker versions of Properties (R1) and (R2). In particular, if \( \lambda \in [0, 1] \) then \( R_{ij}^{(k)}(\lambda) \in [0, 1] \) and each \( R_{ij}^{(k)}(\lambda) \) is complex analytic in \( \lambda \in D_1 \).
By a first step analysis based on the Markov property, we show the following proposition in Section 6.1.

**Proposition 5.3.** The set of functions $R_{i,j}^{(k)}(\lambda)$ introduced in Eq. (5.5) satisfy Eq. (3.2) for $\lambda \in [0, 1]$.

In Section 6.2 we use Perron–Frobenius theory to prove:

**Proposition 5.4.** The Markov chain $\{W_n\}_{n \geq 0}$ is transient.

As a direct corollary, we have:

**Corollary 5.5.** The set of functions $R_{i,j}^{(k)}(\lambda)$ introduced in Eq. (5.5) satisfy $R_{i,j}^{(k)}(1) < 1$.

**Proof.** Proposition 5.4 implies that $R_{i,j}^{(k)}(1) = P_{e_i}(T(0, A_{i,j}^{(k)}) < \infty) < 1$. ■

In Section 6.3 we demonstrate using techniques from the analysis of branching processes that

**Proposition 5.6.** For a given $\lambda \in [0, 1]$, the only solution to Eq. (3.2) satisfying Property (R2) is given by the $R_{i,j}^{(k)}(\lambda)$ introduced in Eq. (5.5).

For $x, y \in X$, define the generating function

$$S(x, y; \lambda) := \sum_{n=0}^{\infty} P_x(W_n = y) \lambda^n. \quad (5.6)$$

Next, define the generating function associated with first visits from an arbitrary word to be

$$R(x, y; \lambda) := \sum_{n=0}^{\infty} P_x(T(0, y) = n) \lambda^n. \quad (5.7)$$

Note that these functions are identically zero unless $x, y \in X_i$ for some $1 \leq i \leq N$ by Lemma 2.5. For any $w \in X_i$, we have $S(e_i, w; \lambda) = R(e_i, w; \lambda) S(w, w; \lambda)$ using the strong Markov property with the first hitting time of $w$. In Section 6.4 by obtaining an exponential bound on $P_{e_i}(W_n = e_i)$, we show the following proposition:

**Proposition 5.7.** The functions $S(x, y; \lambda)$ and $R(x, y; \lambda)$ introduced in Eqs. (5.6) and (5.7), respectively, have radii of convergence strictly greater than 1.

These ingredients allow us to prove Proposition 3.1.

**Proof of Proposition 5.1.** We will show that the $R_{i,j}^{(k)}$ defined in Eq. (5.5) are the unique functions satisfying Properties (R1), (R3).

Since $R_{i,j}^{(k)}(\lambda) = R(e_i, A_{i,j}^{(k)}; \lambda)$ for each $(i,j,k)$, by Proposition 5.7 there is an $\epsilon > 0$ such that $R_{i,j}^{(k)}$ is complex analytic in $D_{1+\epsilon}$ as required for Property (R1). Property (R2) is satisfied for $\lambda \in [0, 1)$ by definition and for $\lambda = 1$ by Corollary 5.5. By Proposition 5.3 and Property (R1), the $R_{i,j}^{(k)}$ satisfy Eq. (3.2) in $D_{1+\epsilon}$ as required for Property (R3). Finally, by Proposition 5.6 and Property (R1), the $R_{i,j}^{(k)}$ are unique which completes the proof. ■
We define
\[ G_i(\lambda, z) := E_i \left[ \sum_{n=0}^{\infty} z^{|W_n| L} \lambda^n \right]. \] (5.8)

Let \( K(\lambda, z) \) be the \( 2N \times 2N \) matrix whose blocks are
\[ K(\lambda, z) := \begin{pmatrix} 0 & B(1; \lambda, z) \\ B(-1; \lambda, z) & 0 \end{pmatrix}, \] (5.9)

where \( B(k; \lambda, z) \) is defined in Eq. (3.3) and \( 0 \) is the zero matrix.

For the proof of Theorem 3.2 we require the following two propositions whose proofs are postponed until Section 6.5.

**Proposition 5.8.** In the region \((\lambda, z) \in (0, 1] \times (0, 1] \), the matrix \( K(\lambda, z) \) is irreducible and it has non-negative entries. In the region \((\lambda, z) \in (0, 1) \times (0, 1] \), the spectral radius of \( K(\lambda, z) \) is strictly less than 1.

**Proposition 5.9.** Let \( s(\lambda) \) and \( v(i) \) be \( N \)-vectors whose entries satisfy \( s_j(\lambda) = S(e_j, e_j; \lambda), v_j(i) = \delta_{ij} \) and let \( \bar{s}(\lambda) \) and \( \bar{v}(i) \) be the \( 2N \)-vectors \( \bar{s}(\lambda) = (s(\lambda), s(\lambda)) \) and \( \bar{v}(i) = (v(i), v(i)) \). Then, we have
\[ G_i(\lambda, z) = S(e_i, e_i; \lambda) + \sum_{d=1}^{\infty} \bar{v}^T(i) K(\lambda, z)^d \bar{s}(\lambda), \] (5.10)

where \( G_i(\lambda, z) \) is defined in Eq. (5.8) and \( K(\lambda, z) \) is defined in Eq. (5.9).

We are now ready to prove Theorem 3.2.

**Proof of Theorem 3.2.** It is enough to show Theorem 3.2 with initial condition \( W_0 = e_i \) for each \( 1 \leq i \leq N \). We will apply Theorem 5.1 with \( Y_n = |W_n| L \). Then, \( G_i \) in Eq. (5.8) is \( \mathcal{G} \) in Eq. (5.1).

By Proposition 5.8 there is an \( \epsilon > 0 \) such that the geometric series in Eq. (5.10) converges for \( \lambda, z \in (1 - \epsilon, 1) \). Therefore, in this region, we have
\[ \sum_{d=0}^{\infty} \bar{v}^T(i) K(\lambda, z)^d \bar{s}(\lambda) = \bar{v}^T(i) (I - K(\lambda, z))^{-1} \bar{s}(\lambda) = \frac{\bar{v}^T(i) \text{adj}[I - K(\lambda, z)] \bar{s}(\lambda)}{\text{det}[I - K(\lambda, z)]}, \] (5.11)

where \( \text{adj}[\cdot] \) is the adjugate matrix. Applying this to Eq. (5.10) gives
\[ G_i(\lambda, z) = \frac{C_i(\lambda, z)}{\text{det}[I - K(\lambda, z)]}, \] (5.12)

where \( C_i(\lambda, z) \) depends on \( \text{adj}[I - K] \) and on \( S(e_i, e_i; \lambda) \). Therefore, we have that \( C_i(\lambda, z) \) and \( \text{det}[I - K(\lambda, z)] \) are both polynomial functions of the \( R^{(k)}_{i,j}(\lambda), S(e_i, e_i; \lambda) \). By Proposition 5.7 these functions are analytic for \( \lambda \in D_{1+\epsilon} \) with a possibly smaller \( \epsilon > 0 \). In addition, \( C_i(\lambda, z) \) and \( \text{det}[I - K(\lambda, z)] \) both depend on \( z \) through a finite number of positive powers and hence they are complex analytic in a neighborhood of \( z = 1 \). It follows that \( C_i(\lambda, z) \) and \( \text{det}[I - K(\lambda, z)] \) are complex analytic in the region \( \lambda \in D_{1+\epsilon}, |z - 1| < \epsilon \).
It remains to show that \( C_i(1, 1) \neq 0 \). By Proposition 5.8, \( K(\lambda, z) \) is non-negative and irreducible so, by Perron–Frobenius theory, there is a simple eigenvalue \( \mu_{PF}(\lambda, z) \) of \( K(\lambda, z) \) equal to the spectral radius of \( K(\lambda, z) \). Also by Proposition 5.8, we have \( \mu_{PF}(\lambda, z) < 1 \) when \((\lambda, z) \in [0, 1) \times (0, 1] \). We can write

\[
G_i(\lambda, 1) = 1 - \frac{1}{1 - \lambda} = \frac{C_i(\lambda, 1)}{\det[I - K(\lambda, 1)]}.
\]

Therefore, we must have \( \mu_{PF}(1, 1) \geq 1 \). By Proposition 5.7, the entries of \( K(\lambda, 1) \) are complex analytic in a neighborhood of \( \lambda = 1 \). Therefore, \( \mu_{PF}(1, 1) \) is also a complex analytic in a neighborhood of \( \lambda = 1 \) and so we have \( \mu_{PF}(1, 1) = 1 \). The characteristic polynomial of \( K(\lambda, 1) \) satisfies

\[
\det[x I - K(\lambda, 1)] = (x - \mu_{PF}(\lambda, 1)) k(x, \lambda),
\]

where \( k(x, \lambda) \) is a polynomial in \( x \) with \( k(\mu_{PF}(\lambda, 1), \lambda) \neq 0 \). This implies that the function \( \det[I - K(\lambda, 1)] \) has a simple zero at \( \lambda = 1 \). Hence, by Eq. (5.13), we have that \( C_i(1, 1) \neq 0 \).

We conclude that Theorem 5.1 applies to the random variables \( \{\vert W_n \vert_L \}_{n \geq 0} \) with \( g(\lambda, z) = \det[I - K(\lambda, z)] \). Referring to Eq. (5.9), this can also be written in the form

\[
g(\lambda, z) = \det[I - B(1; \lambda, z) B(-1; \lambda, z)].
\]

This completes the proof of Theorem 3.2.

6. Proofs of the main steps

The following subsections contain the proofs of the propositions for the proof of Theorem 3.2.

6.1. Proof of Proposition 5.3

Proof of Proposition 5.3 By Lemma 2.5 we have

\[
P(T(m, xy) = \ell \mid W_m = x) = P_{e(x)}(T(0, y) = \ell).
\]

We can assume that \((i, j, k) = (1, 2, 1)\). Conditioning on possible values of \( W_1 \) gives

\[
E_{e_1}[\lambda^{T(0,A_{1,2}^{(1)})} | W_1 = A_{1,2}^{(1)}] = p_{1,2}^{(1)} E_{e_1}[\lambda^{T(0,A_{1,2}^{(1)})} | W_1 = A_{1,2}^{(1)}]
\]

\[
+ \sum_{m \neq 1,2} p_{1,m}^{(1)} E_{e_1}[\lambda^{T(0,A_{1,2}^{(1)})} | W_1 = A_{1,m}^{(1)}]
\]

\[
+ \sum_{m \neq 1} p_{1,m}^{(-1)} E_{e_1}[\lambda^{T(0,A_{1,2}^{(1)})} | W_1 = A_{1,m}^{(-1)}].
\]

Let us consider these terms line-by-line. In line (6.2a), on the event \( W_1 = A_{1,2}^{(1)} \) we have \( T(0, A_{1,2}^{(1)}) = 1 \), thus \( E[\lambda^{T(0,A_{1,2}^{(1)})} | W_1 = A_{1,2}^{(1)}] = \lambda \).
Consider any term in the sum on line (6.2b). On the event \{W_1 = A^{(1)}_{1,m}\} we have
\[T(0, A^{(1)}_{1,2}) = 1 + T(1, A^{(1)}_{m,2})\] by definition. Hence, by Eq. (6.1),
\[E_{e_1}[\lambda^{T(0,A^{(1)}_{1,2})} | W_1 = A^{(1)}_{1,m}] = E_{e_m}[\lambda^{1+T(0,A^{(1)}_{m,2})}] = \lambda R^{(1)}_{m,2}(\lambda).\] (6.3)
Consider any term in the sum on line (6.2c). If the process is at \(A^{(-1)}_{1,m}\) with \(m \neq 1\) and visits
\(A^{(1)}_{1,2}\) at a later time, then by Lemma 2.5 it must visit \(e_1\) before visiting \(A^{(1)}_{1,2}\). On the event
\(\{W_1 = A^{(-1)}_{1,m}\}\) the first visit to \(e_1\) happens at step \(1 + T(1, A^{(-1)}_{m,1})\). Conditioning on the value
of \(T(0, A^{(-1)}_{m,1})\) and applying Eq. (6.1) gives
\[E_{e_1}[\lambda^{T(0,A^{(1)}_{1,2})} | W_1 = A^{(-1)}_{1,m}] = E_{e_m}[\lambda^{1+T(0,A^{(-1)}_{m,1},A^{(1)}_{1,2})}] = \lambda R^{(-1)}_{m,1}(\lambda) R^{(1)}_{1,2}(\lambda).\] (6.4)
Collecting all these terms gives Eq. (3.2).

6.2. Proof of Proposition 5.4.

Proof of Proposition 5.4. Suppose that \(\{W_n\}_{n \geq 0}\) is recurrent for a given initial condition. Then, by Lemma 2.5 it is recurrent for any initial condition. Moreover, for a given initial condition \(e_i\), the hitting times \(T(0, A^{(k)}_{i,j})\) are almost surely finite, hence \(R^{(k)}_{i,j}(1) = 1\) for each
\((i, j, k)\).

Let \(D^{(k)}_{i,j}(\lambda) := dR^{(k)}_{i,j}/d\lambda;\) we have \(0 < D^{(k)}_{i,j}(\lambda) < \infty\) for \(0 < \lambda < 1\) (see Remark 5.2). Define the \(2N(N - 1)\)-vectors \(d := (D^{(k)}_{i,j})\) and \(r := (R^{(k)}_{i,j})\). Differentiating Eq. (3.2) with respect to \(\lambda\) gives the linear system
\[d = \lambda^{-1}r + M(\lambda)d,\] (6.5)
where \(M(\lambda)\) is a \(2N(N - 1) \times 2N(N - 1)\) matrix whose entries are

\[M(\lambda)_{(i',j',k')}^{(i,j,k)} = \begin{cases} \lambda p^{(k)}_{i',i} & i' \neq j, j' = j, k' = k \\ \lambda p^{(-k)}_{i,i'} R^{(k)}_{i,j}(\lambda) & j' = i, k' \neq k \\ \lambda \sum_{k \neq i} p^{(-k)}_{i,k} R^{(-k)}_{k,i}(\lambda) & i = i', j = j', k = k'. \end{cases}\] (6.6)

By Remark 5.2, \(M(\lambda)\) extends continuously to \(\lambda \in [0, 1]\).

By Lemma A.1 below, the matrix \(M(\lambda)\) is primitive for \(\lambda \in (0, 1]\), and hence has Perron–
Frobenius eigenvalue \(\mu(\lambda) > 0\). There is a corresponding eigenvector \(v(\lambda)\) with positive entries, and all other eigenvalues of \(M(\lambda)\) are smaller than \(\mu(\lambda)\) in norm (see for example [30]). Since \(r > 0\), by Eq. (6.5) we have \(d > Md\) and multiplying this vector inequality from the left with \(v^T\) gives
\[v^T d > v^T Md = \mu v^T d.\] (6.7)
This shows that \(\mu(\lambda) < 1\) for \(\lambda \in (0, 1)\) by the positivity of \(d\). We will show that \(\mu(\lambda_0) > 1\) for some \(\lambda_0 \in (0, 1)\) which proves the lemma by contradiction.
We will now show that there is a non-zero vector \( u = (u_{i,j}^{(k)}) \) that satisfies \((M(1) - I)u = 0\) and
\[
u_{i,j}^{(k)} = (1 - \delta_{j,1})\tilde{u}_{j}^{(k)} - (1 - \delta_{i,1})\tilde{u}_{i}^{(k)}, \tag{6.8}\]
for some nonzero numbers \( \{\tilde{u}_{i}^{(k)} : 1 \leq i \leq N, k \in \{-1, 1\}\} \). We introduce the \( N \times 2N \) matrix
\[
\begin{align*}
J &:= (x_{i,(\ell,m)})_{1 \leq i, \ell \leq N, m \in \{-1, 1\}}, \\
x_{i,(\ell,m)} &:= -p_{i,\ell}^{(m)}(1 - \delta_{\ell,1})(1 - \delta_{\ell,i}) + \delta_{i,\ell}(1 - \delta_{\ell,1})\sum_{j \neq i} p_{i,j}^{(m)}. \tag{6.9a}
\end{align*}
\]
Suppose that \( u \) is of the form in Eq. \((6.8)\). Then for each \((i, j, k)\) we have
\[
\begin{align*}
[(M(1) - I)u]_{(i,j,k)} &= \sum_{\ell \neq i} p_{i,\ell}^{(k)} u_{\ell,j}^{(k)} + \sum_{\ell \neq i} p_{i,\ell}^{(-k)} u_{\ell,i}^{(-k)} - u_{i,j}^{(k)} \sum_{j \neq i} p_{i,j}^{(k)} \\
&= \sum_{\ell \neq i} \sum_{m = \pm 1} p_{i,\ell}^{(m)} \left[ (1 - \delta_{i,1})\tilde{u}_{i}^{(m)} - (1 - \delta_{\ell,1})\tilde{u}_{\ell}^{(m)} \right] \\
&= [Ju]_{i,k}, \tag{6.10c}
\end{align*}
\]
where we have used \( R_{i,j}^{(k)}(1) = 1 \) in Eq. \((6.10a)\). Hence \( Ju = 0 \) implies \((M(1) - I)u = 0\) if \( u \) is given by Eq. \((6.8)\). Since \( J \) has dimensions \( N \times 2N \), there is a non-trivial vector \( \tilde{u} \) in its nullspace. Then the vector \( u \) defined via Eq. \((6.8)\) satisfies \((M(1) - I)u = 0\) which shows that 1 is an eigenvalue of \( M(1) \). Since \( u_{i,j}^{(k)} = -u_{j,i}^{(k)} \) for all \( i, j, k \), the entries of the corresponding eigenvector cannot be all positive, so 1 cannot be the Perron–Frobenius eigenvalue of \( M(1) \). This implies \( \mu(1) > 1 \).

Since \( \mu(\lambda) \) is a simple root of the characteristic polynomial of \( M(\lambda) \), it is continuous in the coefficients of that polynomial \[36\]. These coefficients are continuous functions of \( \lambda \) on \([0, 1]\). Hence, \( \mu(\lambda) \) is a continuous function of \( \lambda \), so there exists a \( \lambda_0 \in (0, 1) \) such that \( \mu(\lambda_0) > 1 \), and we have the desired contradiction.

6.3. **Proof of Proposition \[5.6\]** The proof of Proposition \[5.6\] follows well-known techniques in branching processes, see Sevastyanov \[31\].

**Proof of Proposition \[5.6\]** Fix a given \( \lambda \in [0, 1] \). We begin considering Eq. \((3.2)\) which, in order to simplify the notation, we express as a \( 2N(N - 1) \)-vector equation,
\[
q = f(q, \lambda). \tag{6.11}
\]
The vector entries are labeled by the multi-index \( \ell = (i, j, k) \), with
\[
f_{\ell}(q, \lambda) = f_{i,j}^{(k)}(q, \lambda) = \lambda \left( p_{i,j}^{(k)} + \sum_{m \neq i,j} p_{i,m}^{(k)} q_{m,j}^{(-k)} + \sum_{m \neq i} p_{i,m}^{(-k)} q_{m,i}^{(-k)} \right). \tag{6.12}
\]
Observe that each \( f_\ell(q, \lambda) \) is a quadratic polynomial in \( q \) whose coefficients are non-negative numbers summing to at most 1. Therefore, for \( 2N(N-1) \)-vectors \( x, y \), we have

\[
0 \leq x < y \leq 1 \implies 0 < f(x, \lambda) < f(y, \lambda) \leq 1,
\]

where 1 is a vector with unit entries. Let \( \{a_n(\lambda)\}_{n \geq 0} \) be a sequence defined recursively as

\[
a_{n+1}(\lambda) = f(a_n(\lambda), \lambda), \quad a_0(\lambda) = f(0, \lambda).
\]

By Lemma A.2 and Corollary 5.5

\[
(a_n(\lambda))_\ell \leq \sum_{m=0}^{2^{n+1}-1} P_{e_i}(T(0, A_{i,j}^{(k)}) = m) \lambda^m \leq R_{i,j}^{(k)}(\lambda) < 1, \quad n \geq 0.
\]

Let \( q^*(\lambda) \) be a vector whose entries are \( q_i^*(\lambda) = R_{i,j}^{(k)}(\lambda) \). By Eqs. (6.13) and (6.15) we have that \( \{a_n(\lambda)\}_{n \geq 0} \) is a strictly increasing sequence bounded above by \( q^*(\lambda) \). Therefore, \( \lim_{n \to \infty} a_n(\lambda) = q^*(\lambda) \), where \( f(q^*, \lambda) = q^* \) and \( 0 < q^*(\lambda) < 1 \).

We will now show that \( q^*(\lambda) \) is the only solution to Eq. (6.11) satisfying \( 0 < q(\lambda) < 1 \). Suppose that there is an \( r(\lambda) \neq q^*(\lambda) \) such that \( r = f(r, \lambda) \) and \( 0 < r(\lambda) < 1 \). Applying the function \( f(\cdot, \lambda) \) repeatedly to both sides of the inequality \( 0 < r(\lambda) \) and using Eq. (6.12) we get \( q^*(\lambda) \leq r(\lambda) \). We will drop the \( \lambda \) dependence for the remainder of this proof. Draw the line \( z(\theta) = q^* + (r - q^*)\theta \); there will be a point on this line \( \tilde{r} \leq 1 \) such that \( \tilde{r}_{\ell} = 1 \) for some \( \ell \) and \( \tilde{r} > 1 \). We therefore have \( f_\ell(\tilde{r}) \leq 1 = \tilde{r}_{\ell} \) by Eq. (6.13). Let \( \varphi(\theta) = f_\ell(z(\theta)) - z_\ell(\theta) \), then we have

\[
\varphi(0) = f_\ell(q^*) - q^*_\ell = 0, \quad \varphi(\tilde{\theta}) = f_\ell(\tilde{r}) - \tilde{r}_\ell \leq 0.
\]

By direct computation, and since \( f_\ell \) is a quadratic polynomial with non-negative coefficients,

\[
\varphi''(\theta) = \sum_{m,n} (r_m - q^*_m)(r_n - q^*_n) \frac{\partial^2 f_\ell}{\partial z_m \partial z_n} \geq 0.
\]

Since \( f_\ell \) is nonlinear, \( \varphi''(\theta) \) is not identically zero so convexity gives \( \varphi(\theta) < 0 \) for \( \theta \in (0, \tilde{\theta}) \). In particular, \( \varphi(1) < 0 \), so \( f_\ell(r) < r_\ell \): a contradiction. \( \blacksquare \)

6.4. Proof of Proposition 5.7. In this section, we show that the functions \( S(x, y; \lambda) \) and \( R(x, y; \lambda) \) introduced in Eqs. (5.6) and (5.7), respectively, have radii of convergence strictly greater than 1. This would follow from Proposition 8.1 in Lalley (2001) [17], however the proof of this proposition is incomplete. A correction was provided to us by the author via personal communication and he kindly allowed us to reproduce the corrected proof here, adapted to our case.

To prove Proposition 5.7, we require the following Lemma which is a consequence of the Azuma–Hoeffding inequality for bounded submartingales [7].
Lemma 6.1. Let $\xi_1, \xi_2, \ldots$ be a sequence of Bernoulli random variables adapted to a filtration $\{F_n\}_{n \geq 0}$. Assume that there exists $p > 0$ such that, for every $n \geq 1$,

$$P(\xi_{n+1} = 1 | F_n) \geq p. \quad (6.18)$$

Then for every $\alpha < p$, there exist $\beta < 1$ and $C < \infty$ such that $\forall n \geq 1$,

$$P\left(\sum_{i=1}^{m} \xi_i \leq \alpha m \text{ for some } m \geq n\right) \leq C \beta^n. \quad (6.19)$$

Proof of Proposition 5.7. Fix any $1 \leq i \leq N$. For any $x, y \in X_i$ we have $P_x(W_n = y) \geq P_x(T(0, y) = n)$ so it will be sufficient to show that the radius of convergence of $S(x, y; \lambda)$ is strictly greater than 1; we begin by showing that that the radii of convergence of the $S(x, y; \lambda)$ are all the same.

Let $x, x', y, y' \in X_i$. By Lemma 2.5, there exists a positive-probability path from $x$ to $y$ that passes through $x'$ then $y'$ on the way. Suppose that the shortest path from $x$ to $x'$ has $\ell_1$ steps, from $y$ to $y'$ has $\ell_2$ steps, and from $x'$ to $y'$ has $n'$ steps. Let $k = \ell_1 + \ell_2$, then by the Markov property we have that, for all $n \geq n'$,

$$P_x(W_n + k = y) \geq P_x(W_{n+\ell_1+\ell_2} = y, W_{n+\ell_1} = y', W_{\ell_1} = x') \geq \epsilon P_{x'}(W_n = y'), \quad (6.20)$$

where $\epsilon > 0$ is independent of $n$. By the same argument, there exists a $k' \geq 0$ and $\epsilon' > 0$ such that, for sufficiently large $n$,

$$P_{x'}(W_{n+k'} = y') \geq \epsilon' P_x(W_n = y). \quad (6.21)$$

Therefore, $S(x, y; \lambda)$ and $S(x', y'; \lambda)$ have the same radii of convergence.

We will now show that $S(e_i, e_i; \lambda)$ has radius of convergence strictly greater than 1. To do this, we will show that there are constants $C < \infty$ and $\beta < 1$ such that $P_{e_i}(W_n = e_i) \leq C \beta^n$. We will write $e := e_i$ for the remainder of this proof.

By Assumptions 2.2, 2.3 and Lemma 2.5, the quantity

$$c(w) := P_w(|W_n| > |w| \text{ for all } n > 0)$$

takes one of a finite number of possible values (since, if $w$ is non-empty, $c(w)$ only depends on the last generator in the reduced representation of $w$.) By Proposition 5.4, our process is transient and hence $c(A^{(k)}_{i,j}) > 0$ for at least one $(i, j, k)$. For any word $w$, we can append at most four generators at the end of $w$ to produce a word ending in $A^{(k)}_{i,j}$ in a way that the length of the word strictly increases during this process. Hence by the Markov property, there exists a $q > 0$ such that

$$P_w(|W_n| > |w| \text{ for all } n > 0) \geq q \quad \forall w \in X. \quad (6.22)$$
We fix $m \geq 1$ to be the smallest integer such that $mq > 1 - q$. We define the stopping times $\tau_k$ inductively such that $\tau_0 = 0$ and

$$\tau_{k+1} := \min_{n > \tau_k} \{|W_n| - |W_{\tau_k}| \in \{-1, 0, m\}\}. \quad (6.23)$$

By the transience of our process, $\tau_k < \infty$ almost surely and $W_{\tau_k}$ is well-defined.

Consider the event $\{W_n = e\}$ and take $j = \max\{k : \tau_k < n\}$. By Assumptions 2.2 and 2.3, we have $|W_{m+1}| - |W_m| \in \{-1, 0, m\}$ for each $m \geq 0$. It follows that $|W_{\tau_j}| \leq 1$ since otherwise there would be a $\tau_j < n' < n$ such that $|W_{n'}| - |W_{\tau_j}| \in \{-1, 0\}$, so $\tau_{j+1} \leq n' < n$ which is a contradiction. Therefore, on the event $\{W_n = e\}$ we have $|W_{\tau_j}| \leq 1$ since otherwise there would be a $\tau_j < n' < n$ such that $|W_{n'}| - |W_{\tau_j}| \in \{-1, 0\}$, so $\tau_{j+1} \leq n' < n$ which is a contradiction. Therefore, we must have $\tau_j \leq n$.

Let $\gamma \in (0, 1)$, then partitioning $\{W_n = e\}$ using the event $\{\tau_k = n\}$ for some $k \geq \gamma n$ and its complement we get the upper bound

$$P_e(W_n = e) \leq P(\tau_{\lceil \gamma n \rceil} > n) + P_e(W_{\tau_k} = e \text{ for some } k \geq \gamma n), \quad (6.24)$$

where $\lceil \cdot \rceil$ is the ceiling function. We will show that there is a $\gamma$ such that each term in Eq. (6.24) has an exponential bound.

Define the Bernoulli random variables $\{\xi_k\}_{k \geq 1}$ by

$$\xi_k := \begin{cases} 1 & mk < \tau_j \leq m(k + 1) \text{ for some } j \geq 0 \\ 0 & \text{otherwise.} \end{cases} \quad (6.25)$$

These are adapted to the filtration $\mathcal{F}_{m(k+1)}$, $k \geq 1$ where $\mathcal{F}_n$ is the $\sigma$-algebra generated by the first $n$ steps of the process. By Assumption 2.4, we have

$$P_w(|W_1| = |w| + 1) > 0, \quad \forall w \in X. \quad (6.26)$$

By Eq. (6.26) and the Markov property, there exists an $\alpha > 0$ such that

$$P_e(\xi_{k+1} = 1 | \mathcal{F}_{mk}) \geq \alpha, \quad \forall k \geq 1. \quad (6.27)$$

Suppose that $\tau_\ell > mb$ for some $\ell, b \geq 1$. If $\sum_{k=1}^b \xi_k > \ell$, then there would be at least $\ell$ distinct intervals of length $m$ up to time $mb$ containing a $\tau_k$, which would imply that $\tau_\ell \leq mb$. Therefore, we must have

$$P(\tau_\ell > mb) \leq P\left(\sum_{k=1}^b \xi_k \leq \ell\right), \quad (6.28)$$

Consider the first term in Eq. (6.24) and take $\gamma = \alpha/(2m)$. By Eqs. (6.27), (6.28) and Lemma 6.1 we have that there exist constants $C < \infty$ and $0 < \beta < 1$ such that

$$P(\tau_{\lceil \gamma n \rceil} > n) \leq P\left(\sum_{k=1}^{\ell} \xi_k \leq (\alpha/2)\ell \text{ for some } \ell \geq n/m\right) \leq C\beta^n, \quad \forall n \geq 0. \quad (6.29)$$
Next, define the Bernoulli random variables \( \{ \zeta_k \}_{k \geq 1} \) by

\[
\zeta_{k+1} := \begin{cases} 
1 & |W_{\tau_{k+1}} - |W_{\tau_k}| = m \\
0 & |W_{\tau_{k+1}} - |W_{\tau_k}| \in \{-1, 0\}.
\end{cases}
\]  

These are adapted to the filtration \((\mathcal{F}_{\tau_k})_{k \geq 0}\). By Eq. (6.22) and the Markov property, we have

\[
P(\zeta_{k+1} = 1 | \mathcal{F}_{\tau_k}) \geq q, \quad \forall k \geq 0.
\]  

(6.31)

We fix \( r < q \) such that

\[
mr - (1 - r) := \Delta > 0.
\]  

(6.32)

By Eq. (6.31) and Lemma 6.1, there exist constants \( K < \infty \) and \( 0 < \delta < 1 \) such that

\[
P_e(\sum_{\ell=1}^k \zeta_{\ell} \leq rk \text{ for some } k \geq n) \leq K\delta^n, \quad \forall n \geq 0.
\]  

(6.33)

If \( \sum_{\ell=1}^k \zeta_{\ell} \geq rk \), then \( |W_{\tau_k}| \geq mrk - (1 - r)k = k\Delta \) by Eq. (6.32). Therefore,

\[
P_e(|W_{\tau_k}| < k\Delta \text{ for some } k \geq n) \leq K\delta^n \quad \forall n \geq 0.
\]  

(6.34)

Considering the second term in Eq. (6.24), by Eq. (6.34) we have that for all \( n \geq 0 \),

\[
P_e(W_{\tau_k} = e \text{ for some } k \geq \gamma n) \leq P_e(|W_{\tau_k}| < k\Delta \text{ for some } k \geq \gamma n) \leq K(\delta^n)^n,
\]  

(6.35)

where \( 0 < \delta^\gamma < 1 \). Combining Eqs. (6.29) and (6.35) with Eq. (6.24) gives that \( P_e(W_n = e) \) has an exponential bound, which in turn implies that the radius of convergence of \( S(e, e; \lambda) \) is strictly greater than 1. Therefore, we also have that the radius of convergence of \( S(x, y; \lambda) \) is strictly greater than 1 for all \( x, y \in X_i \). Since \( i \) was arbitrary, we conclude that the radius of convergence of \( S(x, y; \lambda) \) is strictly greater than 1 for all \( x, y \in X \), as required. 

\[\blacksquare\]

### 6.5. Proof of Proposition 5.8 and Proposition 5.9

We begin by establishing two lemmas we will need for the proofs of Propositions 5.8 and 5.9.

**Lemma 6.2.** Suppose that \( w_1, w_2 \in X_i \) for some \( 1 \leq i \leq N \) and \( w_1 \neq w_2 \). Consider the reduced composition

\[
w_1^{-1}w_2 = \prod_{\ell=1}^d A_{i, j, \ell}^{(k_{j, \ell})} \quad d \geq 1
\]  

(6.36)

given by Lemma 2.1. Then, we have

\[
\mathcal{R}(w_1, w_2; \lambda) = \prod_{i=1}^d R_{i, j, \ell}^{(k_{j, \ell})}(\lambda).
\]  

(6.37)

**Proof.** This follows from Lemma 2.5 and the strong Markov property applied at the stopping times when the process reaches the intermediate words \( \prod_{i=1}^m A_{i, j, \ell}^{(k_{j, \ell})} \) for each \( 1 \leq m \leq d \). 

\[\blacksquare\]
We introduce the set of words
\[ F_{i,j}^{(k)}(d) := \{ w \in X_i : |w| = d, \ t(w) = j, \ \mathcal{R}(w) = k \}, \tag{6.38} \]
where \( \mathcal{R}(w) = k \) if \(|w| \geq 1 \) and the first letter of \( w \) in its reduced representation is of the form \( A_{i,j}^{(k)} \). A representative member of \( F_{i,j}^{(k)}(d) \) is shown in Figure 6. Let \( B(d, k; \lambda, z) \) be the matrix defined by the product
\[ B(d, k; \lambda, z) := \prod_{i=1}^{d} B((-1)^{i+1}k; \lambda, z), \tag{6.39} \]
where \( B(k; \lambda, z) \) is defined in Eq. (3.3). Then, we have the following lemma.

**Lemma 6.3.** The entries of \( B(d, k; \lambda, z) \) satisfy
\[ B_{i,j}(d, k; \lambda, z) = \sum_{w \in F_{i,j}^{(k)}(d)} z^{|w|} \mathcal{R}(e_i, w; \lambda), \quad d \geq 1. \tag{6.40} \]

**Proof.** The proof is by induction on \( d \). The set \( F_{i,j}^{(k)}(1) \) only contains the word \( A_{i,j}^{(k)} \). We therefore have
\[ B_{i,j}(1, k; \lambda, z) = B_{i,j}(k; \lambda, z) = (1 - \delta_{i,j})z^{|A_{i,j}^{(k)}|} R_{i,j}^{(k)}(\lambda) = \sum_{w \in F_{i,j}^{(k)}(1)} z^{|w|} \mathcal{R}(e_i, w; \lambda). \tag{6.41} \]
Let \( k_d := (-1)^{d+1}k \). By the induction hypothesis and Lemma 6.2,
\[ B_{i,j}(d + 1, k; \lambda, z) = \sum_{\ell} B_{i,\ell}(d, k_d; \lambda, z) B_{\ell,j}(-k_d; \lambda, z) \]
\[ = \sum_{\ell} \sum_{w \in F_{i,\ell}^{(k)}(d)} z^{|w|} \mathcal{R}(e_i, w; \lambda)(1 - \delta_{\ell,j})z^{|A_{\ell,j}^{(-k_d)}|} R_{\ell,j}^{(-k_d)}(\lambda) \]
\[ = \sum_{\ell} (1 - \delta_{\ell,j}) \sum_{w \in F_{i,\ell}^{(k)}(d)} z^{|w| + |A_{\ell,j}^{(-k_d)}|} \mathcal{R}(e_i, w A_{\ell,j}^{(-k_d)}; \lambda) \]
\[ = \sum_{w \in F_{i,j}^{(k)}(d+1)} z^{|w|} \mathcal{R}(e_i, w; \lambda), \tag{6.42} \]
where we have used Eq. (2.5) in Eq. (6.42).

We are now ready to prove Propositions 5.8 and 5.9.

**Proof of Proposition 5.8.** In the region \((\lambda, z) \in (0, 1] \times (0, 1]\), the non-negativity of \(K(\lambda, z)\) immediately follows from the definitions of \(| \cdot |_L\) and \(R_{k,i,j}(\lambda)\). For the remainder of the proof, consider the region \((\lambda, z) \in (0, 1) \times (0, 1]\). Observe that, by Eq. (6.39), we have

\[
K_{2d} = \begin{pmatrix} B_{2d,1;\lambda,z} & 0 \\ 0 & B_{2d,-1;\lambda,z} \end{pmatrix} \quad d \geq 1
\]

(6.43a)

\[
K_{2d-1} = \begin{pmatrix} 0 & B_{2d-1,1;\lambda,z} \\ B_{2d-1,-1;\lambda,z} & 0 \end{pmatrix} \quad d \geq 1.
\]

(6.43b)

The matrix \(B(k;\lambda,z)\) has zeroes on its diagonal and strictly positive entries elsewhere; hence it is primitive. It follows that \(B_{d,k,i,j}(\lambda,z)\) has strictly positive entries for \(d \geq 2\), so \(K(\lambda, z)\) is irreducible with period 2. Next, define the double generating function

\[
F_{i,j}^{(k)}(\lambda, z) := \sum_{n=0}^{\infty} E_{e_i}[z^n|L| \mathbb{I}(t(W_n) = j, \mathcal{R}(W_n) = k)]\lambda^n
\]

(6.44)

\[
= \sum_{d=1}^{\infty} \sum_{w \in \mathcal{F}_{i,j}^{(k)}} z^{|w|L} S(e_i, w; \lambda),
\]

(6.45)

where \(\mathbb{I}(\cdot)\) is the indicator function. Writing \(S(x,y;\lambda) = \mathcal{R}(x,y;\lambda)S(y,y;\lambda)\) and noting that \(S(y,y;\lambda) \geq 1\), we have that Lemma 6.3 implies

\[
B_{i,j}(d,k;\lambda,z) \leq \sum_{w \in \mathcal{F}_{i,j}^{(k)}} z^{|w|L} S(e_i, w; \lambda) \quad d \geq 1.
\]

(6.46)

Combining Eqs. (6.45) and (6.46) and using the fact that \(\lambda \in (0, 1)\) gives

\[
\sum_{d=1}^{\infty} |K(\lambda, z)|_{\ell,m} \leq \max_{i,j,k} \sum_{d=1}^{\infty} B_{i,j}(d,k;\lambda,z) \leq \max_{i,j,k} F_{i,j}^{(k)}(\lambda,z) \leq \frac{1}{1-\lambda} < \infty,
\]

(6.47)

which shows that the spectral radius of \(K(\lambda, z)\) is strictly less than 1.

**Proof of Proposition 5.9.** We have

\[
G_i(\lambda, z) = \sum_{w \in \mathcal{X}_i} z^{|w|L} S(e_i, w; \lambda) = S(e_i,e_i;\lambda) + \sum_{d=1}^{\infty} \sum_{k,j} \sum_{w \in \mathcal{F}_{i,j}^{(k)}} z^{|w|L} S(e_i, w; \lambda).
\]

(6.48)
Applying Lemma 6.3 gives, for \( d \geq 1, \)

\[
\bar{v}^T(i) K(\lambda, z)^d s(\lambda) = \sum_k v^T(i) B(d, k; \lambda, z) s(\lambda)
\]

\[
= \sum_k \sum_{\ell,j} v_\ell(i) \overline{B}_{\ell,j}(d, k; \lambda, z) s_j(\lambda)
\]

\[
= \sum_k \sum_{\ell,j} \sum_{w \in \mathcal{F}_{\ell,j}(d)} \delta_{i,\ell} z^{\lfloor |w| \rfloor} R(\ell, w; \lambda) S(e_j, e_j; \lambda)
\]

\[
= \sum_k \sum_{\ell,j} \sum_{w \in \mathcal{F}_{\ell,j}(d)} z^{\lfloor |w| \rfloor} S(e_i, w; \lambda).
\]

(6.49)

Substituting Eq. (6.49) into Eq. (6.48) completes the proof.

**Appendix A. Supplemental proofs**

**Proof of Lemma 2.1.** Suppose first that two different reduced words \( w, w' \) are in the same equivalence class. Then there is a sequence of words \( w_0 = w, w_1, w_2, \ldots, w_j = w' \) with \( j \geq 1 \) such that each \( w_i \) is obtained from \( w_{i-1} \) by a single operation, namely an application of Eq. (2.3). We will call such an application an “up” (“down”) type move if it increases (decreases) the number of generating elements in the word.

We start by showing that each local sequence of the form \( w_{i-1} \xrightarrow{\text{up}} w_i \xrightarrow{\text{down}} w_{i+1} \) can be edited to a new sequence \( w_{i-1}, w_{i-1,1}, w_{i-1,2} \ldots w_{i-1,\ell}, w_{i+1} \) such that all down moves appear to the left of all up moves. If the down move does not include the generators that have been inserted by the up move, then the up and down moves commute so we can place the down move first. There are finitely many cases when the down move includes one of the generators created by the up move. One can check the cases one-by-one to ensure that they can be edited as required; we omit the details here.

The previous statement implies that the entire sequence \( w_0, w_1, w_2, \ldots w_j \) can be edited to a new sequence \( \bar{w}_0 = w, \bar{w}_1, \bar{w}_2, \ldots \bar{w}_k = w' \) such that all down moves appear to the left of all up moves, and \( k \geq 1 \). But this cannot happen: since \( w, w' \) are both reduced, \( \bar{w}_0 \rightarrow \bar{w}_1 \) cannot be a down move and \( \bar{w}_{k-1} \rightarrow \bar{w}_k \) cannot be an up move. This shows that an arrow cannot have two different reduced representations.

To finish the proof we need to show that each arrow has a reduced representation. For a given composition of finitely many generating elements, we can use the greedy algorithm to apply down steps until we cannot do it anymore. The resulting word cannot include a product of the form \( A^{(k)}_{i,j} A^{(k)}_{j,k} \), which means that it must be empty or of the form Eq. (2.4). ■

**Proof of Lemma 2.5.** The first statement follows directly from the structure of transition probability function (2.7) and Assumptions 2.2 and 2.3.

To prove (ii), by (i) we may assume that \( w_1 = e_i \) for some \( 1 \leq i \leq N \). By Assumptions (2.2) and 2.3 we have that \( s(W_n) = s(W_{n+1}) \) for all \( n \geq 0 \) so we must have \( s(w_2) = i \) if \( e_i \) and \( w_2 \)}
Lemma A.1. The matrix $M(\lambda)$ defined in Eq. (6.6) is primitive for all $N \geq 3$ and $\lambda \in (0, 1]$.

Proof. To show that $M$ is primitive it is sufficient to show that there is an $n$ such that each entry of $M^n$ is strictly positive. Let $\Gamma(Z)$ be the matrix obtained by replacing all the strictly positive entries of any matrix $Z$ with 1. By the assumption that $\lambda \in (0, 1]$, and by direct computation, we have

$$
\Gamma(M_{(i,j,k)}^{\text{(i,j,k)}}) = \delta_{j,j'}\delta_{k,k'} + \delta_{j',i}(1 - \delta_{k,k'}),
$$

(A.2a)

$$
\Gamma(M_{(i,j,k)}^{\text{(i,j,k')}}) = \delta_{j,j'}\delta_{k,k'} + (1 - \delta_{j,j'})(1 - \delta_{k,k'}) + \delta_{j',i}(1 - \delta_{k,k'}) + (1 - \delta_{j',i})\delta_{k,k'},
$$

(A.2b)

$$
\Gamma(M_{(i,j,k)}^{\text{(i,j',k')}}) = 1.
$$

(A.2c)

We remark that $N \geq 3$ is used between Eqs. (A.2a) and (A.2b).

Lemma A.2. For the sequence $\{a_n(\lambda)\}_{n \geq 0}$ defined in Eq. (6.14), we have

$$
(a_n(\lambda))_\ell \leq \sum_{m=0}^{2^{n+1} - 1} P_{e_i}(T(0, A_{i,j}^{(k)}) = m) \lambda^m \quad n \geq 0.
$$

(A.3)

where $T(m, x)$ is defined in Eq. (5.4) and $\ell = (i, j, k)$.

Proof. The proof is by induction on $n \geq 0$. We will write $t_{i,j}^{(k)}(\alpha) := P_{e_i}(T(0, A_{i,j}^{(k)}) = \alpha)$ to simplify the notation. The base case $n = 0$ is immediate since $t_{i,j}^{(k)}(1) = p_{i,j}^{(k)}$ so $(a_0(\lambda))_\ell = \sum_{m=0}^{2^{n+1} - 1} P_{e_i}(T(0, A_{i,j}^{(k)}) = m) \lambda^m \quad n \geq 0.$
By the induction hypothesis and Eq. (6.13), we have

\[(a_{n+1}(\lambda))_i \leq \lambda p_{i,j}^{(k)} \]

\[+ \sum_{\alpha=0}^{2n+1-1} \sum_{m \neq i,j} P_{i,m} t_{m,j}^{(k)}(\alpha) \lambda^{\alpha+1} \]

\[+ \sum_{\alpha=0}^{2n+1-1} \sum_{\beta=0}^{2n+1-1} P_{i,m} t_{m,i}^{(-k)}(\alpha) t_{i,j}^{(k)}(\beta) \lambda^{\alpha+\beta+1}. \]

(A.4c)

We show that the right side of Eq. (A.4) can be bounded above by

\[\sum_{m=0}^{2n+2-1} P_{e_i} \left( T(0, A_{i,j}^{(k)}) = m \right) \lambda^m. \]

(A.5)

The term in (A.4a) is bounded by

\[\sum_{m=0}^{2n+2-1} P_{e_i} \left( T(0, A_{i,j}^{(k)}) = m, W_1 = A_{i,j}^{(k)} \right) \lambda^m, \]

(A.6)

and the term in line (A.4b) is bounded by

\[\sum_{\alpha=0}^{2n+2-1} \sum_{m \neq i,j} P_{e_i} \left( T(0, A_{i,j}^{(k)}) = \alpha, W_1 = A_{i,m}^{(k)} \right) \lambda^\alpha. \]

(A.7)

Finally, we consider line (A.4c). By Lemma 2.5, we can write

\[P_{e_m} \left( T(0, A_{m,i}^{(-k)} A_{i,j}^{(k)}) = \beta \right) = \sum_{\alpha=0}^{\beta} t_{m,i}^{(-k)}(\alpha) t_{i,j}^{(k)}(\beta - \alpha) \]

(A.8)

with \(t_{i,j}^{(k)}(0) = 0\). For non-negative numbers \(c_{\alpha,\beta}\) indexed by \(0 \leq \alpha \leq n, 0 \leq \beta \leq n\), we have

\[\sum_{\alpha=0}^{n} \sum_{\beta=0}^{n} c_{\alpha,\beta} \leq \sum_{r=0}^{2n} \sum_{\alpha+\beta=r} \sum_{\beta=0}^{\beta} c_{\alpha,\beta}. \]

(A.9)

Therefore, applying Eq. (A.9) then Eq. (A.8) results in

\[\sum_{\alpha=0}^{2n+1-1} \sum_{\beta=0}^{2n+1-1} t_{m,i}^{(-k)}(\alpha) t_{i,j}^{(k)}(\beta) \lambda^{\alpha+\beta} \leq \sum_{r=0}^{2n+2-2} P_{e_m} \left( T(0, A_{m,i}^{(-k)} A_{i,j}^{(k)}) = r \right) \lambda^r. \]

(A.10)

Hence, the quantity on line (A.4c) can be bounded above by

\[\sum_{\alpha=0}^{2n+2-1} \sum_{m \neq i,j} P_{e_i} \left( T(0, A_{i,j}^{(k)}) = \alpha, W_1 = A_{i,m}^{(-k)} \right) \lambda^\alpha, \]

(A.11)

which completes the induction step and hence the proof.
Acknowledgments

The authors are grateful to Steven Lalley for valuable discussions and for his assistance with the proof of Proposition [5.4]. B.V. was partially supported by the NSF award DMS-1712551.

References

[1] C. Bélisle, Windings of random walks, Ann. Prob., 17 (1989), pp. 1377–1402.
[2] C. Bélisle and J. Faraway, Winding angle and maximum winding angle of the two-dimensional random walk, J. Appl. Prob., 28 (1991), pp. 717–726.
[3] M. A. Berger, The random walk winding number problem: convergence to a diffusion process with excluded area, J. Phys. A, 20 (1987), pp. 5949–5960.
[4] M. A. Berger and P. H. Roberts, On the winding number problem with finite steps, Adv. Appl. Prob., 20 (1988), pp. 261–274.
[5] R. Brown, Topology and Groupoids, www.groupoids.org.uk, 2006.
[6] N. Desenonges, On the dynamics of Riccati foliations with nonparabolic monodromy representations, Conformal Geometry and Dynamics of the American Mathematical Society, 23 (2019), pp. 164–188.
[7] D. A. Freedman, On tail probabilities for martingales, Ann. Probab., 3 (1975), pp. 100–118.
[8] X. Geng and G. Iyer, Long time asymptotics of heat kernels and Brownian winding numbers on manifolds with boundary, Apr. 2018. Preprint, https://arxiv.org/abs/1804.00368.
[9] L. A. Gilch, Rate of Escape of Random Walks on Regular Languages and Free Products by Amalgamation of Finite Groups, Discrete Mathematics & Theoretical Computer Science, DMTCS Proceedings vol. AI, Fifth Colloquium on Mathematics and Computer Science (2008), pp. 405–420.
[10] A. Grosberg and H. Frisch, Winding angle distribution for planar random walk, polymer ring entangled with an obstacle, and all that: Spitzer–Edwards–Prager–Frisch model revisited, J. Phys. A, 36 (2003), pp. 8955–8981.
[11] J.-C. Gruet, On the length of the homotopic Brownian word in the thrice punctured sphere, Probability Theory and Related Fields, 111 (1998), pp. 489–516.
[12] Y. Guivarc’h, Sur la loi des grands nombres et le rayon spectral d’une marche aléatoire, Astérisque, 74 (1980), pp. 47–98.
[13] D. Holcman and Z. Schuss, The narrow escape problem, SIAM Review, 56 (2014), pp. 213–257.
[14] K. Itô and H. P. McKean, Diffusion processes and their sample paths, Springer, Berlin, 1965.
[15] M. Kac, W. L. Murdock, and G. Szegő, On the eigen-values of certain hermitian forms, Journal of Rational Mechanics and Analysis, 2 (1953), pp. 767–800.
[16] T. Kato, Perturbation theory for linear operators, vol. 132, Springer Science & Business Media, 2013.
[17] S. Lalley, Random walks on regular languages and algebraic systems of generating functions, Contemp. Math., 287 (2001).
[18] S. P. Lalley, Finite range random walk on free groups and homogeneous trees, Ann. Probab., 21 (1993), pp. 2087–2130.
[19] T. J. Lyons and H. P. McKean, Winding of the plane Brownian motion, Adv. Math., 51 (1984), pp. 212–225.
[20] H. P. McKean, Stochastic Integrals, Academic Press, New York, 1969.
[21] H. P. McKean and D. Sullivan, Brownian motion and harmonic functions on the class surface of the thrice punctured sphere, Adv. Math., 51 (1984), pp. 203–211.
[22] P. Messulam and M. Yor, *On D. Williams’ pinching method and some applications*, Journal of the London Mathematical Society, s2-26 (1982), pp. 348–364.

[23] S. K. Nechaev, *Statistics of Knots and Entangled Random Walks*, World Scientific, Singapore; London, 1996.

[24] J. W. Pitman and M. Yor, *The asymptotic joint distribution of windings of planar Brownian motion*, Bull. Am. Math. Soc., 10 (1984), pp. 109–111.

[25] ———, *Asymptotic laws of planar Brownian motion*, Ann. Probab., 14 (1986), pp. 733–779.

[26] ———, *Further asymptotic laws of planar Brownian motion*, Ann. Probab., 17 (1989), pp. 965–1011.

[27] J. Rudnick and Y. Hu, *The winding angle distribution of an ordinary random walk*, Journal of Physics A: Mathematical and General, 20 (1987), pp. 4421–4438.

[28] ———, *Winding angle of a self-avoiding random walk*, Phys. Rev. Lett., 60 (1988), pp. 712–715.

[29] S. Sawyer and T. Steger, *The rate of escape for anisotropic random walks in a tree*, Probability Theory and Related Fields, 76 (1987), pp. 207–230.

[30] E. Seneta, *Non-negative Matrices and Markov Chains*, Springer Series in Statistics, Springer New York, 2006.

[31] B. A. Sevastyanov, *Branching processes*, Nauka, 1971. (In Russian).

[32] F. Spitzer, *Some theorems concerning 2-dimensional Brownian motion*, Trans. Amer. Math. Soc., 87 (1958), pp. 187–197.

[33] F. Tanaka, *Topological classification of Brownian orbits*, J. Chem. Phys., 137 (2012), p. 104907.

[34] S. Watanabe, *Asymptotic windings of Brownian motion paths on Riemann surfaces*, Acta Applicandae Mathematicae, 63 (2000), pp. 441–464.

[35] H. Wen and J.-L. Thiffeault, *Winding of a Brownian particle around a point vortex*, Philos. Trans. Royal Soc. A, 377 (2019), p. 20180347.

[36] M. Zedek, *Continuity and location of zeros of linear combinations of polynomials*, Proceedings of the American Mathematical Society, 16 (1965), pp. 78–84.