Asymptotic Properties of Random Walks via Generating Functions

Lorenz A. Gilch

Habilitationsschrift
zur Erlangung der Lehrbefugnis (venia docendi)
für das Fach Mathematik

Eingereicht an der
Fakultät für Mathematik, Physik und Geodäsie
Institut für Diskrete Mathematik der
Technischen Universität Graz

Nachstehende Arbeiten werden als Habilitationsschrift eingereicht:
The following research articles are provided as part of the habilitation thesis:

1. **Asymptotic Entropy of Random Walks on Free Products**, *Electronic Journal of Probability*, Volume 16, 76–105, 2011.
2. **Phase Transitions for Random Walk Asymptotics on Free Products of Groups.** *Random Structures and Algorithms*, Vol. 40, Issue 2, 150–181, 2012. (with E. Candellero)
3. **Branching Random Walks on Free Products of Groups.** *Proceedings of the London Mathematical Society*, (3) 104, no. 6, 1085–1120, 2012. (with E. Candellero and S. Müller)
4. **Asymptotic Entropy of Random Walks on Regular Languages over a Finite Alphabet**, accepted for publication (modulo minor revision) in *Electronic Journal of Probability*, 2015.

Contents

Introduction

This thesis is devoted to the study of asymptotic properties of transient random walks $(X_n)_{n\in\mathbb{N}_0}$ in different interrelated context. This includes the investigation of the asymptotic behaviour of return transition probabilities as well as the derivation of several limit theorems, which we will describe below. For this purpose, we apply different mathematical techniques from probability theory (random walks), structure theory (algebra, geometry, and graph theory) and analysis (potential theory). The main technique in our proofs consists of a strong use of generating functions, which are power series in one or several variables, where the coefficients are of particular interest for the underlying problem under consideration. These coefficients are often specific probabilities but may also be some other quantities whose asymptotic behaviour we want to study.

In the following we want to give an informal outline of this thesis by describing the different questions which are investigated. The corresponding research articles can be found in the appendix of this thesis and we refer to them as Publications A, B, C and D.

First, we consider nearest neighbour random walks on free products of lattices of the form $\mathbb{Z}^{d_1} * \ldots * \mathbb{Z}^{d_r}$. These randoms walks arise as a convex combination of random walks on the single factors $\mathbb{Z}^{d_i}$. In this setting we will describe the asymptotic behaviour of the n-step return probabilities $\mathbb{P}[X_n = x \mid X_0 = x]$, $x \in \mathbb{Z}^{d_1} * \ldots * \mathbb{Z}^{d_r}$, as n tends to infinity. More generally, we consider free products of finitely generated groups whose Green functions admit an expansion at their radii of convergence with algebraic-logarithmic terms as singular terms up to sufficiently large order. A complete specification of all different asymptotic behaviours of the form $\varrho^n n^{-\lambda} \log^\kappa n$ (here, ϱ denotes the spectral radius of the underlying random walk) is given in Publication B, including an exact description of the phase transitions.

Second, for finitely generated groups $\Gamma_1, \ldots, \Gamma_r$, we will investigate branching random walks on the free product $\Gamma = \Gamma_1 * \ldots * \Gamma_r$. A branching random walk is a growing cloud of particles on the Cayley graph of Γ which has the following evolution: we start with one particle at some given vertex. At each instant of time each particle produces randomly some offspring, and all particle make one step according to an underlying random walk on Γ. In the weak survival phase (that is, if the mean of offspring particles is between 1 and $1/\varrho$, where ϱ is the spectral radius of the underlying random walk on Γ) the particle cloud will vacate each finite set of vertices and the particle cloud moves towards the geometric boundary of the graph. Our aim is to measure the size of the part of the geometric boundary which is hit by the cloud of particles. This size is measured by use of the box-counting dimension (also called Minkowski dimension) and the Hausdorff dimension of the geometric boundary of the whole graph and the corresponding dimensions of the (random) part of the boundary towards which the cloud of particles converges. Publication C studies these dimensions and answers several other related interesting questions, which completes the picture.

Third, we study the asymptotic entropy of random walks on free products of graphs and regular languages, that is, we investigate the limit $-\frac{1}{n}\mathbb{E}[\log \pi_n(X_n)]$ and show that this limit exists, where π_n is the distribution of X_n. Moreover, we are interested in the question whether the asymptotic entropy varies real-analytically in terms of probability measures of constant support. While existence of entropy for random walks on groups follows directly from Kingman's subadditive ergodic theorem, existence for non-group invariant random walks is not guaranteed a priori due to lack of subadditivity. This problem was the starting point for the investigation of the question whether the entropy exists for (non-group invariant) random walks on free products of graphs and on regular languages, and if so to show its real-analytic behaviour. These problems are solved in Publication A for random walks on free products of graphs and in Publication D for random walks on regular languages.

The plan of this thesis is as follows: in Chapter 1 we will give an introduction to random walks and generating functions, and we will motivate their use. In Section 1.3 we give a short introduction to free products of graphs and groups which form the underlying structure of the random walks considered in Publications A, B and C. In particular, we will give an overview on important research articles which deal in a substantial way with generating functions in this context. The structure of Chapters 2, 3 and 4 is as follows: we formulate the main problems, summarize the most important results of Publications A, B, C and D, explain briefly the strong use of generating function techniques in the proofs, and give a detailed outline of important research articles in the corresponding context. Chapter 2 states the main results of Publication B about the asymptotic behaviour of return probabilities for random walks on free products of lattices and groups, while Chapter 3 gives an overview on branching random walks and the above mentioned related questions concerning boundary dimensions (Publication C). Chapter 4 gives an introduction to asymptotic entropy and presents the main results for random walks on free products of graphs (Publication A) and on regular languages (Publication D).

Chapter 1

Random Walks, Generating Functions and Free Products

1.1 Random Walks

A *random walk* is a time-homogeneous Markov chain $(X_n)_{n\in\mathbb{N}_0}$ on a finite or countable state space $\mathcal{S}$ equipped with a transition matrix $P = \big(p(x,y)\big)_{x,y\in\mathcal{S}}$ (also called transition operator) such that $\mathbb{P}[X_{n+1} = y | X_n = x] = p(x,y)$ for all $x, y \in \mathcal{S}$ and $n \in \mathbb{N}_0$. Recall that a random walk is called *transient* if the random walk returns to any starting point $x \in \mathcal{S}$ after finite time with a probability strictly less than 1 and the random walk is called *irreducible* if, for all $x, y \in \mathcal{S}$, the random walk starting at x can visit y after finite time with positive probability. Throughout this thesis we will consider transient and (in most cases) irreducible random walks only. In particular, we assume $\mathcal{S}$ to be infinite. Recall that a *random walk on a group* Γ is given by a probability measure μ on Γ such that $p(x,y) = \mu(x^{-1}y)$ for all $x, y \in \Gamma$.

Random walks are of big structure theoretical relevance: from the probabilistic point of view one considers random walks which are adapted to different algebraic or geometric structures and one wants to investigate the impact of the structure on the random walk's behaviour. Typical questions concern the asymptotic behaviour of return transition probabilities (see Chapter 2) or the asymptotic speed or asymptotic entropy (see Chapter 4). Vice versa, from the geometric point of view random walks can be used for investigation of the geometric structure of the underlying state space. For a better visualisation, we may always think of random walks on graphs. For more information on this interplay, we refer to Woess [57], which serves as a basic reference throughout this thesis and the attached Publications A-D in the appendix.

1.2 Generating Functions

In this section we motivate the use of *generating functions* which play an important role in asymptotic analysis and probability. They form a powerful tool for the investigation of different problems from various fields of mathematics. This thesis is devoted to applications of generating function techniques to problems related to the asymptotic behaviour of random walks. In the context of random walks generating functions are often power series whose coefficients are some specific transition probabilities. In the following we want to give a short introduction to some important generating functions and motivate their use. For this purpose, we use for the n-step transition probabilities of $(X_n)_{n\in\mathbb{N}_0}$ the notion $p^{(n)}(x,y) := \mathbb{P}[X_n = y \mid X_0 = x]$ for $x, y \in \mathcal{S}$. One of the generating functions of main interest is the well-known *Green function* which is defined as

$$G(x,y|z) := \sum_{n\geq 0} p^{(n)}(x,y)\, z^n, \quad z \in \mathbb{C}.$$

If $(X_n)_{n\in\mathbb{N}_0}$ is irreducible then the Green functions $G(\cdot,\cdot|z)$ have a common radius of convergence $R \geq 1$. Typical questions concern the asymptotic behaviour of the transition probabilities $p^{(n)}(x,y)$ as $n \to \infty$. In many cases one gets a power law of the form $p^{(n\delta)}(x,x) \sim C\varrho^{n\delta} n^d$, $d \in \mathbb{R}$, $C > 0$, where $\varrho := \limsup_{n\to\infty} p^{(n)}(x,x)^{1/n} = 1/R$ is the *spectral radius* of the transition operator P of $(X_n)_{n\in\mathbb{N}_0}$ and $\delta := \gcd\{m \in \mathbb{N} \mid p^{(m)}(x,x) > 0\}$ is the *period* of the random walk (see e.g. [57] for further explanations). This gives a first important motivation for the use of generating functions. In the following chapters we will present more details and applications.

Another important class of generating functions is given by *first visit generating functions* defined as

$$F(x,y|z) := \sum_{n\geq 0} \mathbb{P}[X_n = y, \forall m < n : X_m \neq y \mid X_0 = x]\, z^n,$$

which are closely related to Green functions: by conditioning on the first visit to y we get the following essential equation:

$$G(x,y|z) = F(x,y|z) \cdot G(y,y|z).$$

This interaction between different classes of generating functions is very typical also in other context. In order to analyze some generating function $M(z)$ of interest one uses very often a strategy as follows: one considers "simpler" generating functions $M_1(z), \ldots, M_k(z)$, which are somehow in relation with $M(z)$, and tries to establish some system of equations in the unknown variables $M_1(z), \ldots, M_k(z)$. Then one solves this system and the solutions $M_1(z), \ldots, M_k(z)$ allow a better description of $M(z)$ or even give rise to a formula for $M(z)$, from which one can deduce the aimed results. In particular, this strategy is used in different ways in all the Publications A-D in the appendix. We also refer to the next section where we explain in more detail this interrelation of the involved generating functions in the setting of free products. At this point let us mention the survey article of Woess [58] which outlines the use of generating functions in the study of random walks.

We also want to point out the link between generating functions and boundary theory. The *Martin boundary* of random walks is defined via Martin kernels of the form $K(x,y) := G(x,y|1)/G(o,x|1)$, where $o \in \mathcal{S}$ is some reference point. In order to define the elements of the Martin boundary one lets y tend to infinity and considers the pointwise limits (in the variable x) of the Martin kernels. For more details on Martin boundary we refer e.g. to Woess [55, 57].

Applications of generating function techniques are not necessarily restricted to generating functions with some transition probabilities as coefficients. They also play an important role in combinatorics when one wants to deduce the asymptotic behaviour of the growth of some quantities. For instance, let $(m_n)_{n\in\mathbb{N}}$ be a sequence of some quantities of interest and define the generating function $N(z) = \sum_{n\geq 1} m_n z^n$. Then – if the underlying structure allows this – one can follow the same strategy and analysis as explained above in order to deduce the asymptotic behaviour of m_n as $n \to \infty$.

Furthermore, sometimes it is convenient to deal with *double generating functions*, which are power series in two variables $y, z \in \mathbb{C}$ of the form $V(y,z) = \sum_{m,n\geq 0} v_{m,n} y^m z^n$. In particular, in Publications A and C we use double generating functions.

1.3 Free Products of Graphs

In this section we give a brief introduction to free products of graphs which form the underlying structure of the random walks considered in Publications A, B and C. Free products play an important role in graph and group theory. Stalling's Splitting Theorem states that the Cayley graph of a finitely generated group Γ has more than one (geometric) end if and only if Γ admits a non-trivial decomposition as a free product by amalgamation or an HNN-extension over a finite subgroup. Free products are special cases of amalgamated free products (see e.g. Gilch [24] for definition and examples) and still allow calculations in many situations, while these calculations are getting much more difficult on one-ended non-amenable graphs.

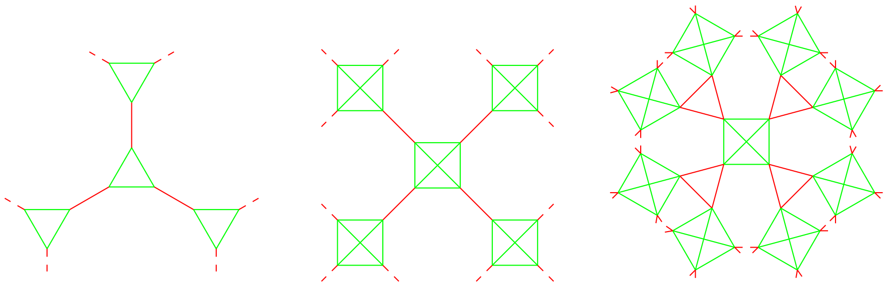
We want to recall the definition of free products. Let $2 \leq r \in \mathbb{N}$ and suppose we are given rooted graphs $G_i = (V_i, o_i)$ for each $i \in \{1,\dots,r\}$, where V_i is the vertex set of G_i and o_i some distinguished vertex. We call o_i the *root* of G_i and we write $\sim_i$ for the adjacency relation on G_i. We set $V_i^\times := V_i \setminus \{o_i\}$. Define

$$V := \left\{ x_1 \dots x_n \,\middle|\, n \in \mathbb{N}, x_i \in \bigcup_{l=1}^{r} V_l^\times, x_j \in V_m^\times \Rightarrow x_{j+1} \notin V_m^\times \right\} \cup \{o\}, \tag{1.1}$$

which is the set of all finite words with letters in $\bigcup_{l=1}^r V_l^\times$ such that no two consecutive letters come from the same $V_i^\times$ and where o denotes the empty word. We have a partial composition law on V: if $w_1 = x_1 \dots x_m, w_2 = y_1 \dots y_n \in V$ with $x_m \in V_i^\times$ and $y_1 \notin V_i^\times$ then the concatenation $w_1 w_2$ is again an element of V. Additionally, we set $w_1 o_j := w_1$ for $j \neq i$, $o_i w_2 := w_2$ and $wo := w =: ow$.

We now define a natural adjacency relation $\sim$ on V as follows: if $u_i \sim_i v_i$ for $u_i, v_i \in V_i$ and if $w \in V$ does not end with a letter in $V_i^\times$, then $wu_i \sim wv_i$.

The *free product* of $G_1, \dots, G_r$ is now given by $G := G_1 * \dots * G_r := (V, o)$ together with the adjacency relation $\sim$. For better visualisation, we explain its graph structure: take copies of $G_1, \dots, G_r$ and glue them together at their roots o_i, which becomes o; inductively, at each vertex $w = x_1 \dots x_m$ with $x_m \in V_i^\times$ attach at w copies of the graphs G_j, $j \neq i$, where each o_j is identified with w. This leads to a cactus-like structure of G, see Figure 1.1.



(a) $K_2 * K_3$ (b) $K_2 * K_4$ (c) $K_3 * K_4$

Figure 1.1: Examples for free products of complete graphs K_n.

Suppose now we are given random walks on G_i governed by transition matrices $P_i = \bigl(p_i(x_i, y_i)\bigr)_{x_i, y_i \in V_i}$, where the entries $p_i(x_i, y_i)$ denote the single-step transition probabilities. We now lift the random walks P_i on the graphs G_i to random walks on G governed by transition matrices $\bar{P}_i = \bigl(\bar{p}_i(u, v)\bigr)_{u,v \in V}$: if $u_i, v_i \in V_i$ and $w \in V$ such that the last letter of w does not belong to V_i, then we set $\bar{p}_i(wu_i, wv_i) := p_i(u_i, v_i)$. Let be $\alpha_1, \dots, \alpha_r \in (0,1)$ with $\sum_{i=1}^r \alpha_i = 1$. A natural way to define a random walk on G is then given by defining the convex combination $P := \alpha_1 \bar{P}_1 + \dots + \alpha_r \bar{P}_r$, which becomes the transition operator of the random walk on G.

A special case is the free product of *groups*: let $\Gamma_1, \dots, \Gamma_r$ be finitely generated groups. Then a (typical) Cayley graph of $\Gamma = \Gamma_1 * \dots * \Gamma_r$ is just the free product of the Cayley graphs of the single groups Γ_i with respect to some given (symmetric) generating sets of the Γ_i's. The random walk on Γ_i is then given by a distribution μ_i on Γ_i such that $p_i(x_i, y_i) = \mu_i(x_i^{-1} y_i)$ for $x_i, y_i \in \Gamma_i$.

The cactus-like structure of free products allows to deduce information about generating functions associated with the free product from the corresponding generating functions on the single factors G_i. The techniques which we use for rewriting probability generating functions on free products in terms of the corresponding generating functions on the single factors of the free product were introduced independently and simultaneously by Cartwright and Soardi [9], Woess [56], Voiculescu [53], and McLaughlin [47]. In the following we want to describe this technique in more detail.

Denote by $G_i(u_i, v_i|z)$ the Green functions on G_i with $u_i, v_i \in V_i$ and we write $G(u, v|z)$, $u, v \in V$, for the Green functions associated with the random walk on G governed by P.

Analogously, we write $F(u,v|z)$ and $F_i(u_i,v_i|z)$ for the first visit generating functions on the free product and on the single factors G_i. A crucial fact is now that one can rewrite some Green functions on G in terms of Green functions on G_i: in [56] a formula for the Green function on G is derived by solving a system of algebraic equations. In particular, this leads to the following important interrelation: for each $i \in \{1,\dots,r\}$, there is a function $\zeta_i(z)$, $z \in \mathbb{C}$, such that for all $u_i, v_i \in V_i$ and each $w \in V$ with last letter not in $V_i^\times$

$$F(wu_i, wv_i|z) = F_i\big(u_i, v_i\big|\zeta_i(z)\big); \tag{1.2}$$

see also Woess [57, Proposition 9.18c]. Furthermore, we have

$$\alpha_i z G(wu_i, wv_i|z) = G_i\big(u_i, v_i\big|\zeta_i(z)\big)\zeta_i(z); \tag{1.3}$$

see [57, Equation (9.20)]. The last equation establishes an important link between Green functions of the free product and Green functions on the single factors. This in turn allows us to deduce information for the asymptotic behaviour of the random walk on G from the asymptotic behaviour of the random walks on the single factors G_i.

Finally, we want to give an overview on some research articles which are closely related to this interplay of generating functions on the free product and on the single factors. Lalley [39] investigated infinite algebraic systems of generating functions of *infinite* free products of finite graphs, where he derived local limit theorems analogous to [56]. For random walks on free products by amalgamation over a finite normal subgroup Cartwright and Soardi [9] derived a formula for the Green function on the amalgamated free product in terms of Green functions on the single factors, which is essentially the same as in Woess [56]. In Gilch [23] different formulas for the rate of escape (also called drift or speed) were derived by strong use of generating function techniques. We will give further literature background in the following chapters.

Chapter 2

Asymptotics of Return Probabilities

The analysis of asymptotic properties of irreducible random walks involves, in particular, the interesting question of the behaviour of n-step return probabilities $p^{(n)}(x,x)$ as $n \to \infty$. Recall that the spectral radius $\varrho = \lim_{n\to\infty} p^{(n\delta)}(x,x)^{1/n\delta} = 1/R$ describes the exponential asymptotic of the return probabilities, where R is the common radius of the Green functions $G(\cdot,\cdot|z)$ and δ the period of the random walk; see e.g. [57]. Now one is interested in describing the asymptotic behaviour of $p^{(n\delta)}(x,x)$ even more precisely by determining the leading sub-exponential term: in many cases one gets that

$$p^{(n\delta)}(x,x) \sim C \cdot \varrho^{n\delta} \cdot n^{\alpha} \tag{2.1}$$

for some suitable constants $C > 0$ and $\alpha \in \mathbb{R}$. For symmetric random walks on groups, Gerl [20] conjectured that the power α is a group invariant. Cartwright's astonishing result [8] disproved this conjecture by giving an example of two symmetric random walks on the free product $\mathbb{Z}^d * \mathbb{Z}^d$ with $d \geq 5$, where each random walk satisfies a law as in (2.1) but with different exponents α. This led to the question of L. Saloff-Coste whether the range of different asymptotic behaviour could still be wider than in the cases considered by Cartwright. This is related with the work of Chatterji, Pittet and Saloff-Coste [11].

Let us summarize some results about the asymptotic behaviour of return transition probabilities related to free products and similar structures. Work in this direction has been done since the 1970's by, amongst others, Gerl, Sawyer, Woess, Cartwright, Soardi and Lalley, see e.g. [9, 21, 37, 50, 56]. For finite range random walks on free groups, the n-step return probabilities behave asymptotically like $C\varrho^n n^{-3/2}$ with $\varrho < 1$; see [21, 37]. In Gerl [20] and Woess [54, 56] free products of finite groups are considered, where finite range random walks obey also a $\varrho^n n^{-3/2}$-law. Picardello and Woess [49] derived the asymptotic behaviour of n-step transition probabilities of random walks on amalgamated free products of compact groups. Lalley [38] calculated the asymptotics for random walks on strings in dependence on positive-recurrence, null-recurrence and transience.

Saloff-Coste's question was the starting point for the investigation of the asymptotic behaviour of return probabilities of nearest neighbour random walks on free products of the form $\mathbb{Z}^{d_1} * \ldots * \mathbb{Z}^{d_r}$, where $r \geq 2$ and $d_1, \ldots, d_r \in \mathbb{N}$. More generally, we studied arbitrary irreducible random walks on free products of groups $\Gamma = \Gamma_1 * \ldots * \Gamma_r$, where the Green functions $G_i(x_i, y_i|z)$ on the single factors Γ_i admit a decomposition of the form

$$G_i(x_i, y_i|z) = f_i(z) + g_i(z) \cdot (R_i - z)^{q_i} \log^{k_i}(R_i - z) \tag{2.2}$$

in a neighbourhood of $z = R_i$, where R_i is the radius of convergence of $G_i(\cdot, \cdot|z)$ and where $f_i(z)$ and $g_i(z)$ are analytic functions with $g(R_i) \neq 0$, $q_i \in \mathbb{R}$ and $k_i \in \mathbb{N}_0$. Here, we call $(R_i - z)^{q_i} \log^{k_i}(R_i - z)$ the *leading singular term*. We note that [B] allows still more general decompositions; see [B, Equation (2.2)]. A decomposition as in (2.2) is given for nearest neighbour random walks on $\mathbb{Z}^{d_i}$; see [57] for simple random walk on $\mathbb{Z}^{d_i}$ and [B, Proposition 6.1] for arbitrary nearest neighbour random walks. In [B] it is shown that there are only up to $r+1$ different asymptotic types for the return probabilities of random walks on free products, which gives a complete answer to Saloff-Coste's question in the case of free products of lattices:

Theorem 2.1. *(see [B, Theorem 1.1])*

Let $2 \leq r \in \mathbb{N}$ and $d_1, \ldots, d_r \in \mathbb{N}$. For each $i \in \{1, \ldots, r\}$, consider a probability measure μ_i on $\mathbb{Z}^{d_i}$ with $\mathrm{supp}(\mu_i) = \{\pm e_j^{(i)} \mid 1 \leq j \leq d_i\}$, where $e_j^{(i)}$ is the j-th unit vector in $\mathbb{Z}^{d_i}$. For any $\alpha_1, \ldots, \alpha_r > 0$ with $\sum_{i=1}^r \alpha_i = 1$, let $\mu := \sum_{i=1}^r \alpha_i \mu_i$ govern an (irreducible) random walk on the free product $\mathbb{Z}^{d_1} * \ldots * \mathbb{Z}^{d_r}$ starting at e, where e denotes the identity of the free product. Denote by ϱ the spectral radius of the random walk governed by μ.

Then the return probabilities $p^{(2n)}(e,e)$ behave asymptotically either like $C \cdot \varrho^{2n} \cdot n^{-d_i/2}$ for $i \in \{1, \ldots, r\}$ or like $C \cdot \varrho^{2n} \cdot n^{-3/2}$ for some constant $C = C_\mu$ depending on μ. Moreover, if all exponents d_i are different and $\min\{d_1, \ldots, d_r\} \geq 5$ then exactly $r+1$ different asymptotic behaviours may occur by choosing the random walk adequately.

Let us note that, for random walks on $\mathbb{Z}^d * \mathbb{Z}^d$ with $d \geq 5$, [8] gave examples for the two possible behaviours of type $n^{-d/2}$ and $n^{-3/2}$. In [57, Proposition 17.13] Cartwright's result is explained that simple random walk obeys a $\varrho^{2n} n^{-d/2}$ law; compare also with Cartwright and Soardi [10].

We even investigated more general laws of the form $C\varrho^{-n\delta} n^{-\lambda} \log^{\kappa} n$ (with ϱ being the spectral radius), which led to the following theorem:

Theorem 2.2. *(see [B, Theorems 3.1 and 4.1])*

Let Γ_1, Γ_2 be finitely generated groups equipped with random walk transition operators P_1 and P_2, whose Green functions $G_1(x_1, y_1|z)$ and $G_2(x_2, y_2|z)$ admit a decomposition of the form (2.2) or as in [B, (2.2)]. Let be $\alpha \in (0,1)$ and define the transition operator $P := \alpha \overline{P}_1 + (1-\alpha)\overline{P}_2$ on $\Gamma_1 * \Gamma_2$, where ϱ denotes its spectral radius and δ its period. Then one of the following different asymptotic behaviours must hold for the random walk on the free product $\Gamma_1 * \Gamma_2$ governed by P:

$$C\varrho^{\delta n} n^{-\lambda_1} \log^{\kappa_1} n \quad \text{or} \quad C\varrho^{\delta n} n^{-\lambda_2} \log^{\kappa_2} n \quad \text{or} \quad C\varrho^{\delta n} n^{-3/2}.$$

The values of $\lambda_1, \lambda_2, \kappa_1, \kappa_2$ can be calculated from the values of q_1, q_2, k_1, k_2 in (2.2).

Moreover, the proof of this theorem gives a complete phase transition analysis: it is precisely formulated where the phase transitions occur and which of the types may occur. Unfortunately, we are not able to present concrete examples with $\kappa_i > 0$; however, the decomposition of the form (2.2) or as in [B, (2.2)] leads also to higher asymptotic orders where the logarithmic term does not vanish; compare with [B, Section 8].

The proof of Theorem 2.2 involves a strong use of generating functions which we sketch in the following. The main technique for deriving these asymptotic behaviours consists of an interaction between the Green functions on $\Gamma_1 * \ldots * \Gamma_r$ and the Green functions on the single factors Γ_i: we plug the expansion given by (2.2) into the formula (1.3) and, by careful analysis, we determine the leading singular algebraic-logarithmic term in the Green function $G(e,e|z)$ of the random walk on the free product. We find out that $G(e,e|z)$ has again a form as in (2.2), where the leading singular term is either inherited from one of the factors $\Gamma_1,\ldots,\Gamma_r$ or which is a square root term.

Having identified the leading singular term of the Green function on the free product we can deduce the asymptotic behaviour of $p^{(n)}(x,x)$ with the help of the well-known *method of Darboux*, which is a standard method for deriving asymptotics. At this point we want to recall Darboux's method. First, the Riemann-Lebesgue Lemma states that if a power series $H(z)=\sum_{n\geq 0} h_n z^n$ has radius of convergence R_H and if $H(z)$ is k-times continuously differentiable on its circle of convergence, then $h_n R_H^n n^k \to 0$ as $n\to\infty$. Thus, one proceeds as follows: identify all singularities on the circle of convergence and subtract parts of the expansion near them such that the remaining part is sufficiently often differentiable on the circle. The asymptotics of the coefficients h_n arise then from the expansions of the leading singular term (and maybe the next higher order singular terms) in the expansion of $H(z)$ in a neighbourhood of $z = R_H$.; we refer to Olver [48, Chap. 8, §9.2] and to Flajolet and Segdewick [17] for the asymptotics of the coefficients in the expansions of standard algebraic-logarithmic singular terms. We remark that another (modern) tool to handle singular expansions is *Singularity Analysis*, which was developed by Flajolet and Odlyzko [16]. However, Darboux's method seems to be more accessible for an application in our setting.

Chapter 3

Asymptotic Behaviour of Branching Random Walks

3.1 Branching Random Walks

Consider a graph G equipped with an irreducible random walk transition operator P and some distribution ν on $\mathbb{N}_0$ with mean $\lambda > 1$. A *branching random walk* (BRW) on G is a growing cloud of particles that move on G in discrete time as follows. The process starts with one single particle at some vertex of G. At each instant of time every particle produces some offspring according to ν and each descendant makes one step in G according to P. Movements and branching of the particles are independent from each other.

Branching random walks are of particular interest at the intersection of abstract mathematics with physics and biology (e.g., bacteria spread out and infect neighbour cells). We give a short overview on the qualitative behaviour of BRWs. Since $\lambda > 1$ the BRW will survive with positive probability; see e.g. Harris [29]. A first natural question is to ask whether the BRW eventually fills up the whole graph, that is, whether every finite set of vertices will eventually be occupied or free of particles. For BRW on Cayley graphs of non-amenable groups one observes the following phase transition, where R denotes the common radius of convergence of the Green functions associated with P: if $\lambda \leq R$ (weak survival phase) then every finite set of vertices will eventually be free of particles (see Benjamini and Peres [5] for $\lambda < R$ and Gantert and Müller [19] for $\lambda = R$); if $\lambda > R$ (strong survival phase) then each vertex will be eventually visited with probability 1 (see [5]). In the first case the trace of the BRW, which consists of all vertices and edges which are visited by the BRW, is a proper subgraph of the Cayley graph of G; see Benjamini and Müller [4].

Further important work in the context of BRW was done by Benjamini and Peres [5, 6], where they give another powerful description of BRWs by tree-indexed random walks. In [4] exponential volume growth of the trace of symmetric (that is, P is symmetric) BRWs on non-amenable Cayley graphs is proved. BRWs are strongly connected with percolation on graphs. For BRWs on free groups and regular trees it turns out that the law of the trace of a BRW is the law of an infinite cluster of some invariant percolation; see Benjamini, Lyons and Schramm [3] and [4].

Finally, let us remark that BRWs have also been studied in the continuous-time setting, see e.g. Harris [29, Chapter V] for Markov branching processes in continuous time and Athreya and Ney [1, Chapter VI] for branching Brownian motion in Euclidean space. Lalley and Sellke [40] studied the phase transition for branching Brownian motion on the hyperbolic disc and Karpelevich, Pechersky and Suhov [35] generalized these results to higher-dimensional Lobachevsky spaces.

3.2 Branching Random Walks on Free Products of Groups

In view of Publication C of this thesis we consider now branching random walks on the Cayley graphs of free products of groups $\Gamma = \Gamma_1 * \ldots * \Gamma_r$, where $\Gamma_1, \ldots, \Gamma_r$ are finitely generated groups. We are interested in the *weak survival phase* $\lambda \in (1, R]$, that is, the cloud of particles moves towards the boundary Ω of the free product and each finite set of vertices will be finally vacated. The limit set Λ of the BRW is the random subset of the boundary Ω that consists of all ends in Ω, where the BRW accumulates. Typical ways of measuring the size of boundaries are by use of the *box-counting dimension* $\mathrm{BD}(M)$ (also known as the *Minkowski dimension*) or the *Hausdorff dimension* $\mathrm{HD}(M)$ for $M \subset \Omega$. Note that existence of the box-counting dimension is not guaranteed a priori. Since the formal definitions of these boundaries and dimensions are not necessary in order to state the main results, we omit an exact definition at this point and refer to [C, page 8].

The starting point for [C] was the work of Hueter and Lalley [31], who studied BRWs on homogeneous trees in the phase $1 < \lambda \leq R$. In particular, they gave a formula for the Hausdorff dimension $\mathrm{HD}(\Lambda)$ and showed that the Hausdorff dimension is at most half of the size of the Hausdorff dimension of Ω. Publication C extends these results to BRWs on free products of groups $\Gamma = \Gamma_1 * \ldots * \Gamma_r$ and free products by amalgamation of finite groups (for a definition, see e.g. [C, Section 3.3]). In order to be able to state the main result we need some definitions: let $F(e, x|z)$ be the first-visit generating function on the free product Γ, where e denotes the identity element of Γ and $x \in \Gamma$. Denote by $|x|$ the word length of x in the sense of the definition in (1.1). Define the double generating function

$$\mathcal{F}_i^+(\lambda|z) := \sum_{x \in \Gamma_i \setminus \{e_i\}} F(e, x|\lambda) z^{|x|} = \sum_{x_i \in \Gamma_i \setminus \{e_i\}} F_i(e_i, x_i|\zeta_i(\lambda)) z^{|x|},$$

where we use (1.2) in the second equation with $F_i(\cdot, \cdot|z)$ being the first visit generating function on Γ_i and e_i being the identity in Γ_i. Now we can formulate our main result:

Theorem 3.1. *(see [C, Theorem 3.5])*
Suppose that ν has a finite second moment. Then the box-counting dimension $\mathrm{BD}(\Lambda)$ *exists and equals the Hausdorff dimension* $\mathrm{HD}(\Lambda)$. *Moreover,* $\mathrm{BD}(\Lambda) = \mathrm{HD}(\Lambda) = -\log z^*/\log 2$, *where z^* is the smallest real positive number with*

$$\sum_{i=1}^{r} \frac{\mathcal{F}_i^+(\lambda|z)}{1+\mathcal{F}_i^+(\lambda|z)} = 1.$$

We remark that the above theorem is a reformulation of [C, Theorem 3.5] with $\alpha = \frac{1}{2}$, where $\alpha \in (0,1)$ is a scaling constant used in the definition of the dimensions. Moreover, an analogous result to Theorem 3.1 is shown for the whole boundary Ω of the free product, see [C, Theorem 3.8]. For the case of BRWs on free products by amalgamation of finite groups we obtain a similar result, see [C, Corollary 3.18].

The behaviour of $\mathrm{HD}(\Lambda)$ when varying λ (in particular, when $\lambda \nearrow R$) is explained in [C, Theorem 3.10], which leads to the following qualitative picture of the mapping $\lambda \mapsto \mathrm{HD}(\Lambda)$:

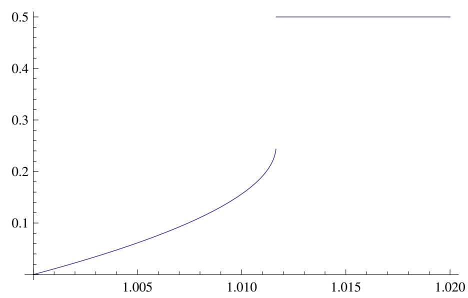


Figure 3.1: Hausdorff dimension $\mathrm{HD}(\Lambda)$ of a BRW on $(\mathbb{Z}/3\mathbb{Z}) * (\mathbb{Z}/2\mathbb{Z})$ in dependence of λ on the x-axis; the discontinuity is at $\lambda = R$, and $\mathrm{HD}(\Lambda) = \mathrm{HD}(\Omega)$ for $\lambda > R$.

In order to establish the formula for $\mathrm{HD}(\Lambda)$ in Theorem 3.1 we make once again a strong use of generating function techniques as explained in Section 1.3: we consider the double generating function $\mathcal{F}(\lambda|z) := \sum_{x\in\Gamma} F(e,x|\lambda) z^{|x|}$ on the free product, which we rewrite in terms of the corresponding double generating functions $\mathcal{F}_i^+(\lambda|z)$ on the single factors:

$$\mathcal{F}(\lambda|z) = \frac{1}{1-\sum_{i=1}^{r} \frac{\mathcal{F}_i^+(\lambda|z)}{1+\mathcal{F}_i^+(\lambda|z)}};$$

see [C, (4.1)]. This equation is the essential key in order to find out the radius of convergence of $\mathcal{F}(\lambda|z)$ (for given λ), from which we can deduce the Hausdorff dimension of Λ. Let us remark that the form of the last equation is characteristic for free products: analogous equations have been established in [23] for the calculation of the asymptotic drift of random walks on free products and in Gilch and Müller [25] for the calculation of the connective constant of self-avoiding walks on free products.

Chapter 4

Asymptotic Entropy of Random Walks

4.1 Random Walks and Asymptotic Entropy

Consider a transient Markov chain $(X_n)_{n\in\mathbb{N}_0}$ on some infinite state space $\mathcal{S}$ and denote by π_n the distribution of X_n. We are interested whether the sequence $-\frac{1}{n}\mathbb{E}[\log \pi_n(X_n)]$ converges, and if so to compute its limit h. If the limit exists it is called the *asymptotic entropy* of $(X_n)_{n\in\mathbb{N}_0}$, which was introduced by Avez [2] for transient random walks on groups. This question is studied for random walks on different structures: Publication A answers this question for random walks on free products of graphs while Publication D gives an answer for random walks on regular languages.

We outline some background on this topic. For random walks on groups, existence of the asymptotic entropy follows from Kingman's subadditive ergodic theorem (see Kingman [36]) due to subadditivity of the sequence $(-\log \pi_n(X_n))_{n\in\mathbb{N}_0}$. In particular, the sequence $-\frac{1}{n}\log \pi_n(X_n)$ converges almost surely to the asymptotic entropy h. However, for non-group invariant random walks existence of the entropy is not guaranteed a priori. In the particular (non-group invariant) cases of free products of graphs and regular languages we have no general subadditivity and only a partial composition law for two words of the free product or regular language; hence, Kingman's theorem can *not* be applied. For more information about entropy of random walks on groups we refer to Kaimanovich and Vershik [32], Derriennic [13] and Kaimanovich and Woess [33].

An important link between drift and harmonic analysis was obtained by Varopoulos [52] who proved that for symmetric finite range random walks on groups the existence of non-trivial bounded harmonic functions is equivalent to a non-zero rate of escape. This result was generalized by Karlsson and Ledrappier [34] to symmetric random walks with finite first moment of the step lengths. This leads to a link between the rate of escape and the entropy of random walks, compare e.g. with [32] and Erschler [14].

Erschler and Kaimanovich [15] asked whether drift and entropy of random walks on groups with finite range vary real-analytically in terms of probability measures of constant support in the following sense: let μ be a finitely supported probability measure on a group Γ_0, whose support $\text{supp}(\mu)$ generates Γ_0 as a semigroup. That is, we can write $\mu = (\mu(s))_{s\in S}$ with some finite $S \subset \Gamma_0$. Then the asymptotic entropy depends on the parameters $(\mu(s))_{s\in S}$ and we ask whether the entropy mapping $\mu \mapsto h = h(\mu)$ is real-analytic. This question can be generalized to non-group invariant random walks, when the transition operator depends on a finite number of parameters only.

The main goals in Publications A and D are to prove that the asymptotic entropy of the underlying random walks exists and to prove its real-analytic behaviour. Compare also with Ledrappier [41], who simultaneously proved this property for finite-range random walks on free groups.

We collect some further recent results on analyticity of drift and entropy. Analyticity of the drift of random walks on free products follows from the formula in Gilch [23]. Ledrappier [42] showed that drift and entropy of finitely supported random walks on hyperbolic groups are Lipschitz. Haïssinksy, Mathieu and Müller [27] showed that the rate of escape of random walks on surface groups varies real-analytically. Mathieu [46] proved that the asymptotic entropy of random walks on nonelementary hyperbolic groups is differentiable. The very recent excellent work of Gouëzel [26] shows that drift and entropy of random walks on hyperbolic groups vary real-analytically. See also Ledrappier and Gilch [22], which is a survey article about regularity of drift and entropy of some random walks.

Finally, we note that the technique of the proofs in Publication A and D was motivated by Benjamini and Peres [6], who showed that, for finite-range random walks on groups, the asymptotic entropy equals the *rate of escape with respect to the Green distance* which is given by $\lim_{n\to\infty} -\frac{1}{n}\log G(e, X_n|1)$; Blachère, Haïssinsky and Mathieu [7] extended the results of [6] to random walks with finite first moment of the step lengths.

4.2 Asymptotic Entropy of Random Walks on Free Products of Graphs

Let $G_1, \ldots, G_r$ be rooted graphs with finite or countable sets of vertices and edges. We now consider a transient random walk on the free product $G = G_1 * \ldots * G_r$. We assume that the random walk is *uniformly irreducible* (for a definition, see e.g. [A, (2.1)]. Denote by R the common radius of convergence of the Green functions $G(\cdot, \cdot|z)$ of the random walk on the free product. Then we have:

Theorem 4.1. *(see [A, Theorem 3.8])*
Assume $R > 1$. Then the asymptotic entropy $h = \lim_{n\to\infty} -\frac{1}{n}\mathbb{E}[\log \pi_n(X_n)]$ exists, is strictly positive and equals the rate of escape with respect to the Green distance.

This generalizes the result of [6] and [7] to a non-group invariant setting. The proof of Theorem 4.1 consists of proving existence of the rate of escape with respect to the Green distance and then showing that this limit equals the entropy. Hence, a strong use of generating functions leads to the proposed result by studying the evolution of $G(o, X_n|1)$ as $n \to \infty$.

A formula for the entropy in terms of generating functions on the single factors G_i is given in [A, Theorem 3.8]. Another equivalent formula for h is derived in [A, Corollary 4.2] with the help of double generating functions and an application of a theorem of Sawyer and Steger [51, Theorem 2.2]. For the special case of free products of groups, a third formula is given in [A, Theorem 5.1], from which the real-analytic behaviour of the entropy in terms of probability measures of constant support is derived:

Corollary 4.2. *(see [A, Corollary 5.2])*

For transient finite-range random walks on free products of groups, the asymptotic entropy varies real-analytically in terms of probability measures of constant support.

Observe that the last corollary holds also for free products of infinite groups which are not necessarily hyperbolic. We remark that the formula for the entropy given in Mairesse and Mathéus [43] for random walks on free products of *finite* groups depends also real-analytically on transition probabilities with fixed support. Furthermore, we proved that $h = \liminf_{n\to\infty} -\frac{1}{n} \log \pi_n(X_n)$ almost surely (see [A, Corollary 3.9]) and that $-\frac{1}{n} \log \pi_n(X_n)$ converges also in L_1 to h (see [A, Corollary 3.11]).

4.3 Random Walks on Regular Languages

Let $\mathcal{A}$ be a finite alphabet and denote by $\mathcal{A}^*$ the set of all finite words over the alphabet $\mathcal{A}$, where we write o for the empty word. Furthermore, let $(X_n)_{n\in\mathbb{N}_0}$ be a transient Markov chain on $\mathcal{A}^*$ with $X_0 = o$ such that at each instant of time the last $K \in \mathbb{N}$ letters of the current word may be replaced by $2K$ other letters and the transition probabilities depend only on the last K letters of the current word and the replacing letters. That is, the random walk depends on a finite number of parameters which describe these single-step transitions. Denote by $\mathcal{L} \subseteq \mathcal{A}^*$ the set of all finite words which can be reached from o with positive probability. Then $\mathcal{L}$ forms a *regular language*, that is, whose words are accepted by a finite-state automaton. For further information on regular languages, we refer to Hopcraft and Ullman [30].

Random walks on strings or regular languages have been studied in many cases. The most important ones (in our context), amongst others, are the works of Malyshev [44, 45], Gairat, Malyshev, Menshikov and Pelikh [18] and Lalley [38]. In [18] the Perron-Frobenius theory was applied in order to state criteria for positive-recurrence, null-recurrence and transience. Moreover, Malyshev proved stabilization laws concerning existence of the stationary distribution and speed in the transient case and convergence of conditional distributions in the ergodic case. Yambartsev and Zamyatin [60] proved a stabilization law for random walks on two semi-infinite strings over a finite alphabet. Lalley [38] also studied random walks on regular languages: he used generating functions in order to deduce the asymptotic behaviour of return probabilities $p^{(n)}(o,o)$ in dependence of the recurrence/transience behaviour of the random walk. He showed that the return probabilities must obey one of three possible power laws. In particular, [38] introduced several other generating functions which are also useful for our purpose. The key fact at this point is that one can calculate these generating functions by establishing and solving a system of quadratic equations which interconnects these generating functions. Lalley's work was the starting point for the investigation of existence of the drift in [24], which in turn was the starting point for Publication D. The main result of D is the following theorem:

Theorem 4.3. *(see [D, Theorem 2.5 and Corollary 2.8])*
Consider a transient random walk $(X_n)_{n\in\mathbb{N}_0}$ on a regular language, which satisfies Assumptions 2.1 and 2.4 in [D]. Then the asymptotic entropy h of $(X_n)_{n\in\mathbb{N}_0}$ exists and varies real-analytically in terms of probability measures of constant support. Moreover, the asymptotic entropy equals the rate of escape with respect to the Green distance.

As in the proof of Theorem 4.1, one first proves that the rate of escape with respect to the Green distance exists and then shows that it equals the asymptotic entropy. This involves once again a strong use of generating functions.

A formula for the entropy is given in [D, Theorem 2.5]: it identifies the asymptotic entropy via the Shannon entropy (*i.e.*, the asymptotic entropy of a stationary process in the sense of Shannon; see e.g. Cover and Thomas [12]) of a hidden Markov chain with an underlying ergodic Markov chain. Similar to [A] we have that $h = \liminf_{n\to\infty} -\frac{1}{n}\log \pi_n(X_n)$ almost surely and that $-\frac{1}{n}\log \pi_n(X_n)$ converges in L_1 to h; see [D, Corollary 2.7].

Another main goal of the paper was not only to show existence of the asymptotic entropy and to deduce formulas for it, but also to show that it varies real-analytically. To this end we cut the random walk into pieces with the help of the concept of cones, a concept which is often used when studying random walks on graphs. It turns out that the pieces of the random walk between the final entries into the different cones form an ergodic Markov chain from which we deduce a hidden Markov chain. The big task is to prove that the involved entropy of this hidden Markov chain varies real-analytically in terms of the entries of the transition matrix of the underlying finite Markov chain. This property can be shown with the result of Han and Marcus [28], who showed that the entropy of a hidden Markov chain varies real-analytically if the transition matrix of the underlying finite Markov chain satisfies some properties. The main task in Publication D is to make a tricky, laborious recoding of an involved finite ergodic Markov chain (that is, the stochastic process arising from the random walk on $\mathcal{L}$ by cutting it into pieces with the help of cones) such that the recoded finite Markov chain leads to the same hidden Markov chain in distribution but additonally satisfying the assumptions of the theorem of Han and Marcus. This leads to the following result:

Theorem 4.4. *(see [D, Theorem 2.6])*

Consider a transient random walk on a regular language satisfying Assumptions 2.1 and 2.4 in [D]. Then the asymptotic entropy h varies real-analytically in terms of probability measures of constant support.

Finally, let us remark that the results in Publication D are by no means direct generalizations of Ledrappier [41], who proved the real-analytic behaviour of the entropy for random walks on free groups. First, the approaches are different: while [41] identifies the asymptotic entropy as the boundary entropy, Publication D identifies the asymptotic entropy as the Shannon entropy of a hidden Markov chain. Second, the results in Publication D adapt to the situation of virtually free groups, which are generalisations of free groups, but the range of applications is considerably wider (e.g., context-free graphs); we refer to [D, Section 2.2.3] for further comments. See also Woess [59], where a similar concept of cones has been used independently.

Bibliography

[1] K. Athreya and P. Ney. *Branching Processes.* Berlin Heidelberg New York: Springer, 1972.

[2] A. Avez. Entropie des groupes de type fini. *C.R.Acad.Sci.Paris Sér. A-B*, 275:A1363–A1366, 1972.

[3] I. Benjamini, R. Lyons, and O. Schramm. Unimodular random trees. *Ergodic Theory and Dynamical Systems*, 35(2):359–373, 2015.

[4] I. Benjamini and S. Müller. On the trace of branching random walks. *Groups, Geometry, and Dynamics*, 2(2):231–247, 2012.

[5] I. Benjamini and Y. Peres. Markov chains indexed by trees. *Ann. Probab.*, 22(1):219–243, 1994.

[6] I. Benjamini and Y. Peres. Tree-indexed random walks on groups and first passage percolation. *Probab. Theory Related Fields*, 98(1):91–112, 1994.

[7] S. Blachère, P. Haïssinsky, and P. Mathieu. Asymptotic entropy and Green speed for random walks on countable groups. *Ann. Probab.*, 36(3):1134–1152, 2008.

[8] D. Cartwright. On the asymptotic behaviour of convolution powers of probabilities on discrete groups. *Monatshefte für Mathematik*, 107:287–290, 1989.

[9] D. Cartwright and P. Soardi. Random walks on free products, quotients and amalgams. *Nagoya Math. J.*, 102:163–180, 1986.

[10] D. Cartwright and P. Soardi. A local limit theorem for random walks on the cartesian product of discrete groups. *Boll. Un. Math. Ital. A VII*, 1(1):107–115, 1987.

[11] I. Chatterji, C. Pittet, and L. Saloff-Coste. Connected Lie groups and property RD. *Duke Math. J.*, 137(3):511–536, 2007.

[12] T. Cover and J. Thomas. *Elements of Information Theory.* Wiley & Sons, 2nd edition, 2006.

[13] Y. Derriennic. Quelques applications du théorème ergodique sous-additif. *Astérisque*, 74:183–201, 1980.

[14] A. Erschler. On the asymptotics of drift. *J. of Math. Sciences*, 121(3):2437–2440, 2004.

[15] A. Erschler and V. Kaimanovich. Continuity of asymptotic characteristics for random walks on hyperbolic groups. *Functional Analysis and Its Applications*, 47(2):152–156, 2013.

[16] P. Flajolet and A. Odlyzko. Singularity analysis of generating functions. *SIAM J. Alg. and Discr. Math.*, 3(2):216–240, 1990.

[17] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2007.

[18] A. S. Gairat, V. A. Malyshev, M. V. Menshikov, and K. D. Pelikh. Classification of Markov chains that describe the evolution of random strings. *Uspekhi Mat. Nauk*, 50(2(302)):5–24, 1995.

[19] N. Gantert and S. Müller. The critical branching Markov chain is transient. *Markov Process. and Rel. Fields.*, 12:805–814, 2007.

[20] P. Gerl. A local central limit theorem on some groups. In *The first Pannonian Symposium on Mathematical Statistics*, Lect. Notes Statistics 8, pages 73–82. Springer, 1981.

[21] P. Gerl and W. Woess. Local limits and harmonic functions for nonisotropic random walks on free groups. *Probab. Theory Related Fields*, 71:341–355, 1986.

[22] L. Gilch and F. Ledrappier. Regularity of the drift and entropy of random walks on groups. *Publ. Mat. Urug.*, 14:147–158, 2013.

[23] L. A. Gilch. Rate of escape of random walks on free products. *J. Aust. Math. Soc.*, 83(I):31–54, 2007.

[24] L. A. Gilch. Rate of Escape of Random Walks on Regular Languages and Free Products by Amalgamation of Finite Groups. In U. Roesler, editor, *Fifth Colloquium on Mathematics and Computer Science*, volume AI of *DMTCS Proceedings*, pages 405–420, Kiel, Germany, 2008. Discrete Math. and Theor. Comput. Sci.

[25] L. A. Gilch and S. Müller. Counting self-avoiding walks on free products of graphs. *Preprint, http://arxiv.org/abs/1509.03209*, 2015.

[26] S. Gouëzel. Analyticity of the entropy and the escape rate of random walks in hyperbolic groups. *Preprint*, 2015.

[27] P. Haïssinsky, P. Mathieu, and S. Müller. Renewal theory for random walks on surface groups. To appear in *Ergodic Theory and Dynamcial Systems*, http://arxiv.org/abs/1304.7625v1, 2012.

[28] G. Han and B. Marcus. Analyticity of entropy rate of hidden Markov chains. *IEEE Trans. Inform. Theory*, 52(12):5251–5266, 2006.

[29] T. E. Harris. *The theory of branching processes.* Die Grundlehren der mathematischen Wissenschaften, Bd. 119. Springer-Verlag, Berlin, 1963.

[30] J. Hopcraft and J. Ullman. *Introduction to Automata Theory, Languages, and Computation.* Addison-Wesley, Reading MA, 1979.

[31] I. Hueter and S. Lalley. Anisotropic branching random walks on homogeneous trees. *Probab. Theory Related Fields*, 116:57–88, 2000.

[32] V. Kaimanovich and A. Vershik. Random walks on discrete groups: boundary and entropy. *Ann. Probab.*, 11:457–490, 1983.

[33] V. Kaimanovich and W. Woess. Boundary and entropy of space homogeneous Markov chains. *Ann. Probab.*, 30:323–363, 2002.

[34] A. Karlsson and F. Ledrappier. Linear drift and Poisson boundary for random walks. *Pure Appl. Math. Q.*, 3:1027–1036, 2007.

[35] F. I. Karpelevich, E. A. Pechersky, and Y. M. Suhov. A phase transition for hyperbolic branching processes. *Comm. Math. Phys.*, 195(3):627–642, 1998.

[36] J. Kingman. The ergodic theory of subadditive processes. *J. Roy. Statist. Soc., Ser. B*, 30:499–510, 1968.

[37] S. Lalley. Finite range random walk on free groups and homogeneous trees. *Ann. Probab.*, 21(4):2087–2130, 1993.

[38] S. Lalley. Random walks on regular languages and algebraic systems of generating functions. *Algebraic Methods in Statistics and Probability, Contemp. Math.*, 287(201–230), 2000.

[39] S. Lalley. Random walks on infinite free products and infinite algebraic systems of generating functions. Preprint, 2002.

[40] S. P. Lalley and T. Sellke. Hyperbolic branching Brownian motion. *Probab. Theory Related Fields*, 108(2):171–192, 1997.

[41] F. Ledrappier. Analyticity of the entropy for some random walks. *Groups, Geometry, and Dynamics*, 6:317–333, 2012.

[42] F. Ledrappier. Regularity of the entropy for random walks on hyperbolic groups. *Ann. Probab.*, 41(5):3582–3605, 2013.

[43] J. Mairesse and F. Mathéus. Random walks on free products of cyclic groups. *J. London Math. Soc.*, 75(1):47–66, 2007.

[44] V. Malyshev. Stabilization laws in the evolution of a random string. *Problems Inform. Transmission*, 30:260–274, 1995.

[45] V. Malyshev. Interacting strings of characters. Technical Report 3057, INRIA, 1996.

[46] P. Mathieu. Differentiating the entropy of random walks on hyperbolic groups. *Ann. Probab.*, 43(1):166–187, 2015.

[47] J. McLaughlin. *Random Walks and Convolution Operators on Free Products.* PhD thesis, New York Univ., 1986.

[48] F. Olver. Asymptotics and Special Functions. *Academic Press, San Diego, California*, 1974.

[49] M. Picardello and W. Woess. Random walks on amalgams. *Monatshefte für Mathematik*, 100:21–33, 1985.

[50] S. Sawyer. Isotropic random walks in a tree. *Zeitschrift f. Wahrscheinlichkeitstheorie*, Verw. Geb. 42:279–292, 1978.

[51] S. Sawyer and T. Steger. The rate of escape for anisotropic random walks in a tree. *Probab. Theory Related Fields*, 76:207–230, 1987.

[52] N. T. Varopoulos. Long range estimates for Markov chains. *Bull. Sci. Math.*, 109:225–252, 1985.

[53] D. Voiculescu. Addition of certain non-commuting random variables. *J. Funct. Anal.*, 66:323–346, 1986.

[54] W. Woess. A local limit theorem for random walks on certain discrete groups. In *Probability measures on groups (Oberwolfach, 1981)*, volume 928 of *Lecture Notes in Math.*, pages 467–477. Springer, Berlin, 1982.

[55] W. Woess. A description of the Martin boundary for nearest neighbour random walks on free products. *Probability Measures on Groups*, VII:203–215, 1985.

[56] W. Woess. Nearest neighbour random walks on free products of discrete groups. *Boll. Un. Mat. Ital.*, 5-B:961–982, 1986.

[57] W. Woess. *Random Walks on Infinite Graphs and Groups.* Cambridge University Press, 2000.

[58] W. Woess. Generating function techniques for random walks on graphs, in "Heat Kernels and Analysis on Manifolds, Graphs, and Metric Spaces". *Contemporary Math.*, 338:391–423, 2003.

[59] W. Woess. Context-free pairs of groups. II – Cuts, tree sets, and random walks. *Discr. Math.*, 312:157–173, 2012.

[60] A. Yambartsev and A. Zamyatin. A stabilization law for two semi-infinite interacting strings of characters. *Bull. Braz. Math. Soc., New Series*, 34(3):361–388, 2003.

Publication A

Asymptotic Entropy of Random Walks on Free Products

L. Gilch

Electronic Journal of Probability,
Volume 16, 76–105, 2011.

Electronic Journal of Probability

Vol. 16 (2011), Paper no. 3, pages 76–105.

Journal URL
http://www.math.washington.edu/~ejpecp/

Asymptotic Entropy of Random Walks on Free Products

Lorenz A. Gilch*
Graz University of Technology
Institut für mathematische Strukturtheorie (Math. C)
Steyrergasse 30
A-8010 Graz
Austria
Email: gilch@TUGraz.at
URL: http://www.math.tugraz.at/~gilch/

Abstract

Suppose we are given the free product V of a finite family of finite or countable sets. We consider a transient random walk on the free product arising naturally from a convex combination of random walks on the free factors. We prove the existence of the asymptotic entropy and present three different, equivalent formulas, which are derived by three different techniques. In particular, we will show that the entropy is the rate of escape with respect to the Greenian metric. Moreover, we link asymptotic entropy with the rate of escape and volume growth resulting in two inequalities.

Key words: Random Walks, Free Products, Asymptotic Entropy.

AMS 2000 Subject Classification: Primary 60J10; Secondary: 28D20, 20E06.

Submitted to EJP on June 23, 2010, final version accepted December 3, 2010.

*Research supported by German Research Foundation (DFG) grant GI 746/1-1.

1 Introduction

Suppose we are given a finite family of finite or countable sets $V_1, \ldots, V_r$ with distinguished vertices $o_i \in V_i$ for $i \in \{1, \ldots, r\}$. The free product of the sets V_i is given by $V := V_1 * \ldots * V_r$, the set of all finite words of the form $x_1 \ldots x_n$ such that each letter is an element of $\bigcup_{i=1}^r V_i \setminus \{o_i\}$ and two consecutive letters arise not from the same V_i. We consider a transient Markov chain $(X_n)_{n \in \mathbb{N}_0}$ on V starting at the empty word o, which arises from a convex combination of transition probabilities on the sets V_i. Denote by π_n the distribution of X_n. We are interested in whether the sequence $\mathbb{E}[-\log \pi_n(X_n)]/n$ converges, and if so, to compute this constant. If the limit exists, it is called the *asymptotic entropy*. In this paper, we study this question for random walks on general free products. In particular, we will derive three different formulas for the entropy by using three different techniques.

Let us outline some results about random walks on free products: for free products of finite groups, Mairesse and Mathéus [21] computed an explicit formula for the rate of escape and asymptotic entropy by solving a finite system of polynomial equations. Their result remains valid in the case of free products of infinite groups, but one needs then to solve an infinite system of polynomial equations. Gilch [11] computed two different formulas for the rate of escape with respect to the word length of random walks on free products of graphs by different techniques, and also a third formula for free products of (not necessarily finite) groups. The techniques of [11] are adapted to the present setting. Asymptotic behaviour of return probabilities of random walks on free products has also been studied in many ways; e.g. Gerl and Woess [10], [28], Sawyer [24], Cartwright and Soardi [5], and Lalley [18], Candellero and Gilch [4].

Our proof of existence of the entropy envolves generating functions techniques. The techniques we use for rewriting probability generating functions in terms of functions on the factors of the free product were introduced independently and simultaneously by Cartwright and Soardi [5], Woess [28], Voiculescu [27] and McLaughlin [22]. In particular, we will see that asymptotic entropy is the rate of escape with respect to a distance function in terms of Green functions. While it is well-known by Kingman's subadditive ergodic theorem (see Kingman [17]) that entropy (introduced by Avez [1]) exists for random walks on groups whenever $\mathbb{E}[-\log \pi_1(X_1)] < \infty$, existence for random walks on other structures is not known a priori. We are not able to apply Kingman's theorem in our present setting, since we have no (general) subadditivity and we have only a partial composition law for two elements of the free product. For more details about entropy of random walks on groups we refer to Kaimanovich and Vershik [14] and Derriennic [7].

An important link between drifts and harmonic analysis was obtained by Varopoulos [26]. He proved that for symmetric finite range random walks on groups the existence of non-trivial bounded harmonic functions is equivalent to a non-zero rate of escape. Karlsson and Ledrappier [16] generalized this result to symmetric random walks with finite first moment of the step lengths. This leads to a link between the rate of escape and the entropy of random walks, compare e.g. with Kaimanovich and Vershik [14] and Erschler [8]. Erschler and Kaimanovich [9] asked if drift and entropy of random walks on groups vary continuously on the probability measure, which governs the random walk. We prove real-analyticity of the entropy when varying the probabilty measure of constant support; compare also with the recent work of Ledrappier [19], who simultaneously proved this property for finite-range random walks on free groups.

Apart from the proof of existence of the asymptotic entropy $h = \lim_{n\to\infty} \mathbb{E}[-\log \pi_n(X_n)]/n$ (Theorem 3.7), we will calculate explicit formulas for the entropy (see Theorems 3.7, 3.8, 5.1 and Corollary 4.2) and we will show that the entropy is non-zero. The technique of our proof of existence of the entropy was motivated by Benjamini and Peres [2], where it is shown that for random walks on groups the entropy equals the rate of escape w.r.t. the Greenian distance; compare also with Blachère, Haïssinsky and Mathieu [3]. We are also able to show that, for random walks on free products of graphs, the asymptotic entropy equals just the rate of escape w.r.t. the Greenian distance (see Corollary 3.3 in view of Theorem 3.7). Moreover, we prove convergence in probability and convergence in L_1 (if the non-zero single transition probabilities are bounded away from 0) of the sequence $-\frac{1}{n}\log \pi_n(X_n)$ to h (see Corollary 3.11), and we show also that h can be computed along almost every sample path as the limes inferior of the aforementioned sequence (Corollary 3.9). In the case of random walks on discrete groups, Kingman's subadditive ergodic theorem provides both the almost sure convergence and the convergence in L_1 to the asymptotic entropy; in the case of general free products there is neither a global composition law for elements of the free product nor subadditivity. Thus, in the latter case we have to introduce and investigate new processes. The question of almost sure convergence of $-\frac{1}{n}\log \pi_n(X_n)$ to some constant h, however, remains open. Similar results concerning existence and formulas for the entropy are proved in Gilch and Müller [12] for random walks on directed covers of graphs. The reasoning of our proofs follows the argumentation in [12]: we will show that the entropy equals the rate of escape w.r.t. some special length function, and we deduce the proposed properties analogously. In the present case of free products of graphs, the reasoning is getting more complicated due to the more complex structure of free products in contrast to directed covers, although the main results about existence and convergence types are very similar. We will point out these difficulties and main differences to [12] at the end of Section 3.2. Finally, we will link entropy with the rate of escape and the growth rate of the free product, resulting in two inequalities (Corollary 6.4).

The plan of the paper is as follows: in Section 2 we define the random walk on the free product and the associated generating functions. In Section 3 we prove existence of the asymptotic entropy and give also an explicit formula for it. Another formula is derived in Section 4 with the help of double generating functions and a theorem of Sawyer and Steger [25]. In Section 5 we use another technique to compute a third explicit formula for the entropy of random walks on free products of (not necessarily finite) groups. Section 6 links entropy with the rate of escape and the growth rate of the free product. Sample computations are presented in Section 7.

2 Random Walks on Free Products

2.1 Free Products and Random Walks

Let $\mathcal{I} := \{1,\dots,r\} \subseteq \mathbb{N}$, where $r \geq 2$. For each $i \in \mathcal{I}$, consider a random walk with transition matrix P_i on a finite or countable state space V_i. W.l.o.g. we assume that the sets V_i are pairwise disjoint and we exclude the case $r = 2 = |V_1| = |V_2|$ (see below for further explanation). The corresponding single and n-step transition probabilities are denoted by $p_i(x,y)$ and $p_i^{(n)}(x,y)$, where $x,y \in V_i$. For every $i \in \mathcal{I}$, we select an element o_i of V_i as the "root". To help visualize this, we think of graphs $\mathcal{X}_i$ with vertex sets V_i and roots o_i such that there is an oriented edge $x \to y$ if and only if $p_i(x,y) > 0$. Thus, we have a natural graph metric on the set V_i. Furthermore, we shall assume that for every $i \in \mathcal{I}$ and every $x \in V_i$ there is some $n_x \in \mathbb{N}$ such that $p_i^{(n_x)}(o_i,x) > 0$. For sake of simplicity we assume $p_i(x,x) = 0$ for every $i \in \mathcal{I}$ and $x \in V_i$. Moreover, we assume that the random walks on V_i

are *uniformly irreducible*, that is, there are $\varepsilon_0^{(i)} > 0$ and $K_i \in \mathbb{N}$ such that for all $x, y \in V_i$

$$p_i(x,y) > 0 \quad \Rightarrow \quad p_i^{(k)}(x,y) \geq \varepsilon_0^{(i)} \quad \text{for some } k \leq K_i. \tag{2.1}$$

We set $K := \max_{i \in \mathscr{I}} K_i$ and $\varepsilon_0 := \min_{i \in \mathscr{I}} \varepsilon_0^{(i)}$. For instance, this property is satisfied for nearest neighbour random walks on Cayley graphs of finitely generated groups, which are governed by probability measures on the groups.

Let $V_i^\times := V_i \setminus \{o_i\}$ for every $i \in \mathscr{I}$ and let $V_*^\times := \bigcup_{i \in \mathscr{I}} V_i^\times$. The *free product* is given by

$$\begin{aligned} V &:= V_1 * \ldots * V_r \\ &= \left\{ x_1 x_2 \ldots x_n \,\middle|\, n \in \mathbb{N}, x_j \in V_*^\times, x_j \in V_k^\times \Rightarrow x_{j+1} \notin V_k^\times \right\} \cup \left\{ o \right\}. \end{aligned} \tag{2.2}$$

The elements of V are "words" with letters, also called *blocks*, from the sets $V_i^\times$ such that no two consecutive letters come from the same V_i. The empty word o describes the root of V. If $u = u_1 \ldots u_m \in V$ and $v = v_1 \ldots v_n \in V$ with $u_m \in V_i$ and $v_1 \notin V_i$ then uv stands for their concatenation as words. This is only a partial composition law, which makes defining the asymptotic entropy more complicated than in the case of free products of *groups*. In particular, we set $uo_i := u$ for all $i \in \mathscr{I}$ and $ou := u$. Note that $V_i \subseteq V$ and o_i as a word in V is identified with o. The *block length* of a word $u = u_1 \ldots u_m$ is given by $\|u\| := m$. Additionally, we set $\|o\| := 0$. The *type* $\tau(u)$ of u is defined to be i if $u_m \in V_i^\times$; we set $\tau(o) := 0$. Finally, $\tilde{u}$ denotes the last letter u_m of u. The set V can again be interpreted as the vertex set of a graph $\mathscr{X}$, which is constructed as follows: take copies of $\mathscr{X}_1, \ldots \mathscr{X}_r$ and glue them together at their roots to one single common root, which becomes o; inductively, at each vertex $v_1 \ldots v_k$ with $v_k \in V_i$ attach a copy of every $\mathscr{X}_j$, $j \neq i$, and so on. Thus, we have also a natural graph metric associated to the elements in V.

The next step is the construction of a new Markov chain on the *free product*. For this purpose, we lift P_i to a transition matrix $\bar{P}_i$ on V: if $x \in V$ with $\tau(x) \neq i$ and $v, w \in V_i$, then $\bar{p}_i(xv, xw) := p_i(v, w)$. Otherwise we set $\bar{p}_i(x,y) := 0$. We choose $0 < \alpha_1, \ldots, \alpha_r \in \mathbb{R}$ with $\sum_{i \in \mathscr{I}} \alpha_i = 1$. Then we obtain a new transition matrix on V given by

$$P = \sum_{i \in \mathscr{I}} \alpha_i \bar{P}_i.$$

The random walk on V starting at o, which is governed by P, is described by the sequence of random variables $(X_n)_{n \in \mathbb{N}_0}$. For $x, y \in V$, the associated single and n-step transition probabilities are denoted by $p(x,y)$ and $p^{(n)}(x,y)$. Thus, P governs a nearest neighbour random walk on the graph $\mathscr{X}$, where P arises from a convex combination of the nearest neighbour random walks on the graphs $\mathscr{X}_i$.

Theorem 3.3 in [11] shows existence (including a formula) of a positive number ℓ_0 such that $\ell_0 = \lim_{n\to\infty} \|X_n\|/n$ almost surely. The number ℓ_0 is called the *rate of escape w.r.t. the block length*. Denote by π_n the distribution of X_n. If there is a real number h such that

$$h = \lim_{n \to \infty} \frac{1}{n} \mathbb{E}\big[-\log \pi_n(X_n)\big],$$

then h is called the *asymptotic entropy* of the process $(X_n)_{n \in \mathbb{N}_0}$; we write $\mathbb{N}_0 := \mathbb{N} \setminus \{0\}$. If the sets V_i are groups and the random walks P_i are governed by probability measures μ_i, existence of the asymptotic entropy rate is well-known, and in this case we even have $h = \lim_{n\to\infty} -\frac{1}{n} \log \pi_n(X_n)$ almost surely; see Derriennic [7] and Kaimanovich and Vershik [14]. We prove existence of h in the case of general free products.

2.2 Generating Functions

Our main tool will be the usage of generating functions, which we introduce now. The *Green functions* related to P_i and P are given by

$$G_i(x_i,y_i|z) := \sum_{n\geq 0} p_i^{(n)}(x_i,y_i)\,z^n \quad \text{and} \quad G(x,y|z) := \sum_{n\geq 0} p^{(n)}(x,y)\,z^n,$$

where $z\in\mathbb{C}$, $x_i,y_i\in V_i$ and $x,y\in V$. At this point we make the *basic assumption* that the radius of convergence R of $G(\cdot,\cdot|z)$ is strictly bigger than 1. This implies *transience* of our random walk on V. Thus, we may exclude the case $r=2=|V_1|=|V_2|$, because we get recurrence in this case. For instance, if all P_i govern *reversible* Markov chains, then $R>1$; see [29, Theorem 10.3]. Furthermore, it is easy to see that $R>1$ holds also if there is some $i\in\mathcal{I}$ such that $p_i^{(n)}(o_i,o_i)=0$ for all $n\in\mathbb{N}$.
The *first visit generating functions* related to P_i and P are given by

$$\begin{aligned} F_i(x_i,y_i|z) &:= \sum_{n\geq 0}\mathbb{P}\bigl[Y_n^{(i)}=y_i, \forall m\leq n-1: Y_m^{(i)}\neq y_i \mid Y_0^{(i)}=x_i\bigr]z^n \quad \text{and}\\ F(x,y|z) &:= \sum_{n\geq 0}\mathbb{P}\bigl[X_n=y, \forall m\leq n-1: X_m\neq y \mid X_0=x\bigr]z^n, \end{aligned}$$

where $\bigl(Y_n^{(i)}\bigr)_{n\in\mathbb{N}_0}$ describes a random walk on V_i governed by P_i. The stopping time of the first return to o is defined as $T_o:=\inf\{m\geq 1\mid X_m=o\}$. For $i\in\mathcal{I}$, define

$$\overline{H}_i(z) := \sum_{n\geq 1}\mathbb{P}[T_o=n, X_1\notin V_i^{\times}]\,z^n \quad \text{and} \quad \xi_i(z) := \frac{\alpha_i z}{1-\overline{H}_i(z)}.$$

We write also $\xi_i:=\xi_i(1)$, $\xi_{\min}:=\min_{i\in\mathcal{I}}\xi_i$ and $\xi_{\max}:=\max_{i\in\mathcal{I}}\xi_i$. Observe that $\xi_i<1$; see [11, Lemma 2.3]. We have $F(x_i,y_i|z)=F_i\bigl(x_i,y_i|\xi_i(z)\bigr)$ for all $x_i,y_i\in V_i$; see Woess [29, Prop. 9.18c]. Thus,

$$\xi_i(z) := \frac{\alpha_i z}{1-\sum_{j\in\mathcal{I}\setminus\{i\}}\sum_{s\in V_j}\alpha_j p_j(o_j,s) z F_j\bigl(s,o_j\big|\xi_j(z)\bigr)}.$$

For $x_i\in V_i$ and $x\in V$, define the stopping times $T_{x_i}^{(i)}:=\inf\{m\geq 1\mid Y_m^{(i)}=x_i\}$ and $T_x:=\inf\{m\geq 1\mid X_m=x\}$, which take both values in $\mathbb{N}\cup\{\infty\}$. Then the *last visit generating functions* related to P_i and P are defined as

$$\begin{aligned} L_i(x_i,y_i|z) &:= \sum_{n\geq 0}\mathbb{P}\bigl[Y_n^{(i)}=y_i, T_{x_i}^{(i)}>n \mid Y_0^{(i)}=x_i\bigr]z^n,\\ L(x,y|z) &:= \sum_{n\geq 0}\mathbb{P}\bigl[X_n=y, T_x>n \mid X_0=x\bigr]z^n. \end{aligned}$$

If $x=x_1\dots x_n, y=x_1\dots x_n x_{n+1}\in V$ with $\tau(x_{n+1})=i$ then

$$L(x,y|z) = L_i\bigl(o_i,x_{n+1}\,\big|\,\xi_i(z)\bigr); \tag{2.3}$$

this equation is proved completely analogously to [29, Prop. 9.18c]. If all paths from $x\in V$ to $w\in V$ have to pass through $y\in V$, then

$$L(x,w|z) = L(x,y|z)\cdot L(y,w|z);$$

this can be easily checked by conditioning on the last visit of y when walking from x to w. We have the following important equations, which follow by conditioning on the last visits of x_i and x, the first visits of y_i and y respectively:

$$\begin{aligned} G_i(x_i,y_i|z) &= G_i(x_i,x_i|z)\cdot L_i(x_i,y_i|z) = F_i(x_i,y_i|z)\cdot G_i(y_i,y_i|z), \\ G(x,y|z) &= G(x,x|z)\cdot L(x,y|z) = F(x,y|z)\cdot G(y,y|z). \end{aligned} \tag{2.4}$$

Observe that the generating functions $F(\cdot,\cdot|z)$ and $L(\cdot,\cdot|z)$ have also radii of convergence strictl bigger than 1.

3 The Asymptotic Entropy

3.1 Rate of Escape w.r.t. specific Length Function

In this subsection we prove existence of the rate of escape with respect to a specific length function. From this we will deduce existence and a formula for the asymptotic entropy in the upcoming subsection.

We assign to each element $x_i \in V_i$ the "length"

$$l_i(x_i) := -\log L(o,x_i|1) = -\log L_i(o_i,x_i|\xi_i).$$

We extend it to a length function on V by assigning to $v_1\dots v_n \in V$ the length

$$l(v_1\dots v_n) := \sum_{i=1}^n l_{\tau(v_i)}(v_i) = -\sum_{i=1}^n \log L(o,v_i|1) = -\log L(o,v_1\dots v_n|1).$$

Observe that the lengths can also be negative. E.g., this can be interpreted as height differences. The aim of this subsection is to show existence of a number $\ell \in \mathbb{R}$ such that the quotient $l(X_n)/n$ tends to ℓ almost surely as $n\to\infty$. We call ℓ the *rate of escape w.r.t. the length function $l(\cdot)$*.

We follow now the reasoning of [11, Section 3]. Denote by $X_n^{(k)}$ the projection of X_n to the first k letters. We define the *k-th exit time* as

$$\mathbf{e}_k := \min\bigl\{m\in\mathbb{N}_0 \bigm| \forall n\geq m : X_n^{(k)} \text{ is constant}\bigr\}.$$

Moreover, we define $\mathbf{W}_k := X_{\mathbf{e}_k}$, $\tau_k := \tau(\mathbf{W}_k)$ and $\mathbf{k}(n) := \max\{k\in\mathbb{N}_0 \mid \mathbf{e}_k \leq n\}$. We remark that $\|X_n\|\to\infty$ as $n\to\infty$, and consequently $\mathbf{e}_k<\infty$ almost surely for every $k\in\mathbb{N}$; see [11, Prop. 2.5]. Recall that $\widetilde{\mathbf{W}}_k$ is just the laster letter of the random word $X_{\mathbf{e}_k}$. The process $(\tau_k)_{k\in\mathbb{N}}$ is Markovian and has transition probabilities

$$\hat{q}(i,j) = \frac{\alpha_j}{\alpha_i}\frac{\xi_i}{\xi_j}\frac{1-\xi_j}{1-\xi_i}\Bigl(\frac{1}{(1-\xi_j)G_j(o_j,o_j|\xi_j)}-1\Bigr)$$

for $i\neq j$ and $\hat{q}(i,i)=0$; see [11, Lemma 3.4]. This process is positive recurrent with invariant probability measure

$$\begin{aligned} \nu(i) &= C^{-1}\cdot\frac{\alpha_i(1-\xi_i)}{\xi_i}\bigl(1-(1-\xi_i)G_i(o_i,o_i|\xi_i)\bigr), \\ \text{where } C &:= \sum_{i\in\mathcal{I}}\frac{\alpha_i(1-\xi_i)}{\xi_i}\bigl(1-(1-\xi_i)G_i(o_i,o_i|\xi_i)\bigr); \end{aligned}$$

see [11, Section 3]. Furthermore, the rate of escape w.r.t. the block length exists almost surely and is given by the almost sure constant limit

$$\ell_0 = \lim_{n\to\infty} \frac{\|X_n\|}{n} = \lim_{k\to\infty} \frac{k}{\mathbf{e}_k} = \frac{1}{\sum_{i,j\in\mathcal{I},i\neq j} \nu(i)\,\alpha_j \frac{1-\xi_j}{1-\xi_i}\gamma'_{i,j}(1)}$$

(see [11, Theorem 3.3]), where

$$\gamma_{i,j}(z) := \frac{1}{\alpha_i}\frac{\xi_i(z)}{\xi_j(z)}\Bigl(\frac{1}{\bigl(1-\xi_j(z)\bigr)G_j\bigl(o_j,o_j\big|\xi_j(z)\bigr)} - 1\Bigr).$$

Lemma 3.1. *The process $\bigl(\widetilde{\mathbf{W}}_k,\tau_k\bigr)_{k\in\mathbb{N}}$ is Markovian and has transition probabilities*

$$q\bigl((g,i),(h,j)\bigr) = \begin{cases} \frac{\alpha_j}{\alpha_i}\frac{\xi_i}{\xi_j}\frac{1-\xi_j}{1-\xi_i} L_j(o_j,h|\xi_j), & \text{if } i\neq j,\\ 0, & \text{if } i=j.\end{cases}$$

Furthermore, the process is positive recurrent with invariant probability measure

$$\pi(g,i) = \sum_{j\in\mathcal{I}} \nu(j) q\bigl((*,j),(g,i)\bigr).$$

Remark: Observe that the transition probabilities $q\bigl((g,i),(h,j)\bigr)$ of $\bigl(\widetilde{\mathbf{W}}_k,\tau_k\bigr)_{k\in\mathbb{N}}$ do not depend on g. Therefore, we will write sometimes an asterisk instead of g.

Proof. By [11, Section 3], the process $\bigl(\widetilde{\mathbf{W}}_k,\mathbf{e}_k-\mathbf{e}_{k-1},\tau_k\bigr)_{k\in\mathbb{N}}$ is Markovian and has transition probabilities

$$\tilde{q}\bigl((g,m,i),(h,n,j)\bigr) = \begin{cases} \frac{1-\xi_j}{1-\xi_i}\sum_{s\in V_j} k_i^{(n-1)}(s)p(s,h), & \text{if } i\neq j,\\ 0, & \text{if } i=j,\end{cases}$$

where $k_i^{(n)}(s) := \mathbb{P}\bigl[X_n = s, \forall l \leq n : X_l \notin V_i^\times | X_0 = o\bigr]$ for $s\in V_*^\times \setminus V_i$. Thus, $\bigl(\widetilde{\mathbf{W}_k},\tau_k\bigr)_{k\in\mathbb{N}}$ is also Markovian and has the following transition probabilities if $i\neq j$:

$$\begin{aligned} q\bigl((g,i),(h,j)\bigr) &= \sum_{n\geq 1}\tilde{q}\bigl((g,*,i),(h,n,j)\bigr) = \frac{1-\xi_j}{1-\xi_i}\sum_{s\in V_j}\sum_{n\geq 1} k_i^{(n-1)}(s)p(s,h)\\ &= \frac{1-\xi_j}{1-\xi_i}\sum_{s\in V_j}\frac{L_j(o_j,s|\xi_j)}{1-\bar{H}_i(1)}p(s,h) = \frac{\alpha_j}{\alpha_i}\frac{\xi_i}{\xi_j}\frac{1-\xi_j}{1-\xi_i}L_j(o_j,h|\xi_j).\end{aligned}$$

In the third equality we conditioned on the last visit of o before finally walking from o to s and we remark that $h\in V_j^\times$. A straight-forward computation shows that π is the invariant probability measure of $\left(\widetilde{\mathbf{W}}_k,\tau_k\right)_{k\in\mathbb{N}}$, where we write $\mathscr{A}:=\left\{(g,i)\,\middle|\, i\in\mathscr{I}, g\in V_i^\times\right\}$:

$$\begin{aligned}
\sum_{(g,i)\in\mathscr{A}} \pi(g,i)\cdot q\big((g,i),(h,j)\big) &= \sum_{(g,i)\in\mathscr{A}}\sum_{k\in\mathscr{I}} \nu(k)\cdot q\big((*,k),(g,i)\big)\cdot q\big((*,i),(h,j)\big)\\
&= \sum_{i\in\mathscr{I}} q\big((*,i),(h,j)\big)\sum_{k\in\mathscr{I}}\nu(k)\sum_{g\in V_i^\times} q\big((*,k),(g,i)\big)\\
&= \sum_{i\in\mathscr{I}} q\big((*,i),(h,j)\big)\sum_{k\in\mathscr{I}}\nu(k)\cdot\hat{q}(k,i)\\
&= \sum_{i\in\mathscr{I}} q\big((*,i),(h,j)\big)\cdot\nu(i)=\pi(h,j).
\end{aligned}$$

□

Now we are able to prove the following:

Proposition 3.2. *There is a number $\ell\in\mathbb{R}$ such that*

$$\ell=\lim_{n\to\infty}\frac{l(X_n)}{n}\quad \textit{almost surely}.$$

Proof. Define $h:\mathscr{A}\to\mathbb{R}$ by $h(g,j):=l(g)$. Then $\sum_{\lambda=1}^k h\big(\widetilde{\mathbf{W}}_\lambda,\tau_\lambda\big)=\sum_{\lambda=1}^k l\big(\widetilde{\mathbf{W}}_\lambda\big)=l(\mathbf{W}_k)$. An application of the ergodic theorem for positive recurrent Markov chains yields

$$\frac{l(\mathbf{W}_k)}{k}=\frac{1}{k}\sum_{\lambda=1}^k h\big(\widetilde{\mathbf{W}}_\lambda,\tau_\lambda\big)\xrightarrow{n\to\infty} C_h:=\int h\,d\pi,$$

if the integral on the right hand side exists. We now show that this property holds. Observe that the values $G_j(o_j,g|\xi_j)$ are uniformly bounded from above for all $(g,j)\in\mathscr{A}$:

$$G_j(o_j,g|\xi_j)=\sum_{n\geq 0} p_j^{(n)}(o_j,g)\,\xi_j^n\leq\frac{1}{1-\xi_j}\leq\frac{1}{1-\xi_{\max}}.$$

For $g\in V_*^\times$, denote by $|g|$ the smallest $n\in\mathbb{N}$ such that $p_{\tau(g)}^{(n)}(o_{\tau(g)},g)>0$. Uniform irreducibility of the random walk P_i on V_i implies that there are some $\varepsilon_0>0$ and $K\in\mathbb{N}$ such that for all $j\in\mathscr{I}$, $x_j,y_j\in V_j$ with $p_j(x_j,y_j)>0$ we have $p_j^{(k)}(x_j,y_j)\geq\varepsilon_0$ for some $k\leq K$. Thus, for $(g,j)\in\mathscr{A}$ we have

$$G_j(o_j,g|\xi_j)\geq\varepsilon_0^{|g|}\xi_j^{|g|\cdot K}\geq\left(\varepsilon_0\,\xi_{\min}^K\right)^{|g|}.$$

Observe that the inequality $|g|\cdot\left|\log\big(\varepsilon_0\,\xi_{\min}^K\big)\right|<\log 1/(1-\xi_{\max})$ holds if and only if $|g|<\log(1-\xi_{\max})/\log(\varepsilon_0\,\xi_{\min}^K)$. Define the sets

$$M_1:=\left\{g\in V_*^\times\,\middle|\,|g|\geq\frac{\log(1-\xi_{\max})}{\log(\varepsilon_0\,\xi_{\min}^K)}\right\},\quad M_2:=\left\{g\in V_*^\times\,\middle|\,|g|<\frac{\log(1-\xi_{\max})}{\log(\varepsilon_0\,\xi_{\min}^K)}\right\}.$$

Recall Equation (2.4). We can now prove existence of $\int h\,d\pi$:

$$
\begin{aligned}
\int |h|\,d\pi &= \sum_{(g,j)\in\mathscr{A}} \left|\log L_j(o_j,g|\xi_j)\right|\cdot\pi(g,j)\\
&\leq \sum_{(g,j)\in\mathscr{A}} \left|\log G_j(o_j,g|\xi_j)\right|\cdot\pi(g,j) + \sum_{(g,j)\in\mathscr{A}} \left|\log G_j(o_j,o_j|\xi_j)\right|\cdot\pi(g,j)\\
&\leq \sum_{(g,j)\in\mathscr{A}:g\in M_1} \left|\log G_j(o_j,g|\xi_j)\right|\cdot\pi(g,j)\\
&\quad + \sum_{(g,j)\in\mathscr{A}:g\in M_2} \left|\log G_j(o_j,g|\xi_j)\right|\cdot\pi(g,j) + \max_{j\in\mathscr{I}} \log G_j(o_j,o_j|\xi_j)\\
&\leq \sum_{(g,j)\in\mathscr{A}:g\in M_1} \left|\log(\varepsilon_0\xi_{\min}^K)^{|g|}\right|\cdot\pi(g,j)\\
&\quad + \sum_{(g,j)\in\mathscr{A}:g\in M_2} \left|\log(1-\xi_{\max})\right|\cdot\pi(g,j) + \max_{j\in\mathscr{I}} \log G_j(o_j,o_j|\xi_j)\\
&\leq \sum_{(g,j)\in\mathscr{A}:g\in M_1} \left|\log(\varepsilon_0\xi_{\min}^K)\right|\cdot|g|\cdot\pi(g,j)\\
&\quad + \left|\log(1-\xi_{\max})\right| + \max_{j\in\mathscr{I}} \log G_j(o_j,o_j|\xi_j) < \infty,
\end{aligned}
$$

since $\sum_{(g,j)\in\mathscr{A}} |g|\cdot\pi(g,j) < \infty$, see [11, Proof of Prop. 3.2]. From this follows that $l(\mathbf{W}_k)/k$ tends to C_h almost surely. The next step is to show that

$$
\frac{l(X_n)-l(\mathbf{W}_{\mathbf{k}(n)})}{n} \xrightarrow{n\to\infty} 0 \quad \text{almost surely.} \tag{3.1}
$$

To prove this, assume now that we have the representations $\mathbf{W}_{\mathbf{k}(n)} = g_1g_2\dots g_{\mathbf{k}(n)}$ and $X_n = g_1g_2\dots g_{\mathbf{k}(n)}\dots g_{\|X_n\|}$. Define $M := \max\left\{|\log(\varepsilon_0\,\xi_{\min}^K)|, |\log(1-\xi_{\max})|\right\}$. Then:

$$
\begin{aligned}
\left|l(X_n)-l(\mathbf{W}_{\mathbf{k}(n)})\right| &= \left|-\sum_{i=\mathbf{k}(n)+1}^{\|X_n\|} \log L_{\tau(g_i)}\big(o_{\tau(g_i)}, g_i \mid \xi_{\tau(g_i)}\big)\right|\\
&\leq \sum_{i=\mathbf{k}(n)+1}^{\|X_n\|} \left|\log \frac{G_{\tau(g_i)}\big(o_{\tau(g_i)}, g_i \mid \xi_{\tau(g_i)}\big)}{G_{\tau(g_i)}\big(o_{\tau(g_i)}, o_{\tau(g_i)} \mid \xi_{\tau(g_i)}\big)}\right|\\
&\leq \sum_{i=\mathbf{k}(n)+1:g_i\in M_1}^{\|X_n\|} \left|\log G_{\tau(g_i)}\big(o_{\tau(g_i)}, g_i \mid \xi_{\tau(g_i)}\big)\right|\\
&\quad + \sum_{i=\mathbf{k}(n)+1:g_i\in M_2}^{\|X_n\|} \left|\log G_{\tau(g_i)}\big(o_{\tau(g_i)}, g_i \mid \xi_{\tau(g_i)}\big)\right|\\
&\quad + \big(\|X_n\|-\mathbf{k}(n)\big)\cdot\left|\log(1-\xi_{\max})\right|
\end{aligned}
$$

$$\begin{aligned}
&\leq \sum_{i=\mathbf{k}(n)+1: g_i \in M_1}^{\|X_n\|} \left|\log(\varepsilon_0\, \xi_{\min}^K)^{|g_i|}\right| \\
&\qquad + \sum_{i=\mathbf{k}(n)+1: g_i \in M_2}^{\|X_n\|} \left|\log(1-\xi_{\max})\right| + \left(\|X_n\| - \mathbf{k}(n)\right) \cdot \left|\log(1-\xi_{\max})\right| \\
&\leq \sum_{i=\mathbf{k}(n)+1: g_i \in M_1}^{\|X_n\|} |g_i| \cdot M + \sum_{i=\mathbf{k}(n)+1: g_i \in M_2}^{\|X_n\|} M + \left(\|X_n\| - \mathbf{k}(n)\right) \cdot M \\
&\leq 3 \cdot M \cdot (n - \mathbf{e}_{\mathbf{k}(n)}).
\end{aligned}$$

Dividing the last inequality by n and letting $n \to \infty$ provides analogously to Nagnibeda and Woess [23, Section 5] that $\lim_{n\to\infty}\big(l(X_n) - l(\mathbf{W}_{\mathbf{k}(n)})\big)/n = 0$ almost surely. Recall also that $k/\mathbf{e}_k \to \ell_0$ and $\mathbf{e}_{\mathbf{k}(n)}/n \to 1$ almost surely; compare [23, Proof of Theorem D] and [11, Prop. 3.2, Thm. 3.3]. Now we can conclude:

$$\frac{l(X_n)}{n} = \frac{l(X_n) - l(\mathbf{W}_{\mathbf{k}(n)})}{n} + \frac{l(\mathbf{W}_{\mathbf{k}(n)})}{\mathbf{k}(n)} \frac{\mathbf{k}(n)}{\mathbf{e}_{k(n)}} \frac{\mathbf{e}_{\mathbf{k}(n)}}{n} \xrightarrow{n\to\infty} C_h \cdot \ell_0 \quad \text{almost surely.} \tag{3.2}$$

□

We now compute the constant C_h from the last proposition explicitly:

$$\begin{aligned}
C_h &= \sum_{(g,j)\in\mathscr{A}} l(g) \cdot \sum_{i\in\mathscr{I}} \nu(i) \cdot q\big((*,i),(g,j)\big) \\
&= \sum_{\substack{i,j\in\mathscr{I},\\ i\neq j}} \sum_{g\in V_j^\times} -\log L_j(o_j,g|\xi_j)\, \nu(i) \frac{\alpha_j}{\alpha_i} \frac{\xi_i}{\xi_j} \frac{1-\xi_j}{1-\xi_i} L_j(o_j,g|\xi_j). \qquad (3.3)
\end{aligned}$$

We conclude this subsection with the following observation:

Corollary 3.3. *The rate of escape ℓ is non-negative and it is the rate of escape w.r.t. the Greenian metric, which is given by $d_{Green}(x,y) := -\log F(x,y|1)$. That is,*

$$\ell = \lim_{n\to\infty} -\frac{1}{n} \log F(e,X_n|1) \geq 0.$$

Proof. By (2.4), we get

$$\ell = \lim_{n\to\infty} -\frac{1}{n} \log F(e,X_n|1) - \frac{1}{n} \log G(X_n,X_n|1) + \frac{1}{n} \log G(o,o|1).$$

Since $F(e,X_n|1) \leq 1$ it remains to show that $G(x,x|1)$ is uniformly bounded in $x \in V$: for $v,w \in V$, the *first visit generating function* is defined as

$$U(v,w|z) = \sum_{n\geq 1} \mathbb{P}\big[X_n = w, \forall m \in \{1,\dots,n-1\} : X_m \neq w \mid X_0 = v\big]\, z^n. \tag{3.4}$$

Therefore,

$$G(x,x|z) = \sum_{n\geq 0} U(x,x|z)^n = \frac{1}{1-U(x,x|z)}.$$

Since $U(x,x|z) < 1$ for all $z \in [1,R)$, $U(x,x|0) = 0$ and $U(x,x|z)$ is continuous, stricly increasing and strictly convex, we must have $U(x,x|1) \leq \frac{1}{R}$, that is, $1 \leq G(x,x|1) \leq \left(1-\frac{1}{R}\right)^{-1}$. This finishes the proof. □

3.2 Asymptotic Entropy

In this subsection we will prove that ℓ equals the asymptotic entropy, and we will give explicit formulas for it. The technique of the proof which we will give was motivated by Benjamini and Peres [2], where it is shown that the asymptotic entropy of random walks on discrete groups equals the rate of escape w.r.t. the Greenian distance. The proof follows the same reasoning as in Gilch and Müller [12].

Recall that we made the assumption that the spectral radius of $(X_n)_{n\in\mathbb{N}_0}$ is strictly smaller than 1, that is, the Green function $G(o,o|z)$ has radius of convergence $R > 1$. Moreover, the functions $\xi_i(z)$, $i \in \mathscr{I}$, have radius of convergence bigger than 1. Recall that $\xi_i = \xi_i(1) < 1$ for every $i \in \mathscr{I}$. Thus, we can choose $\varrho \in (1,R)$ such that $\xi_i(\varrho) < 1$ for all $i \in \mathscr{I}$. We now need the following three technical lemmas:

Lemma 3.4. *For all $m,n \in \mathbb{N}_0$,*

$$p^{(m)}(o,X_n) \leq G(o,o|\varrho) \cdot \left(\frac{1}{1-\max_{i\in\mathscr{I}} \xi_i(\varrho)}\right)^n \cdot \varrho^{-m}.$$

Proof. Denote by $\mathscr{C}_\varrho$ the circle with radius ϱ in the complex plane centered at 0. A straightforward computation shows for $m \in \mathbb{N}_0$:

$$\frac{1}{2\pi i}\oint_{\mathscr{C}_\varrho} z^m \frac{dz}{z} = \begin{cases} 1, & \text{if } m = 0,\\ 0, & \text{if } m \neq 0.\end{cases}$$

Let be $x = x_1 \dots x_t \in V$. An application of Fubini's Theorem yields

$$\begin{aligned}\frac{1}{2\pi i}\oint_{\mathscr{C}_\varrho} G(o,x|z)\, z^{-m}\frac{dz}{z} &= \frac{1}{2\pi i}\oint_{\mathscr{C}_\varrho} \sum_{k\geq 0} p^{(k)}(o,x) z^k\, z^{-m}\frac{dz}{z}\\ &= \frac{1}{2\pi i}\sum_{k\geq 0} p^{(k)}(o,x)\oint_{\mathscr{C}_\varrho} z^{k-m}\frac{dz}{z} = p^{(m)}(o,x).\end{aligned}$$

Since $G(o,x|z)$ is analytic on $\mathscr{C}_\varrho$, we have $|G(o,x|z)| \leq G(o,x|\varrho)$ for all $|z| = \varrho$. Thus,

$$p^{(m)}(o,x) \leq \frac{1}{2\pi}\cdot \varrho^{-m-1}\cdot G(o,x|\varrho)\cdot 2\pi\varrho = G(o,x|\varrho)\cdot\varrho^{-m}.$$

Iterated applications of equations (2.3) and (2.4) provide

$$G(o,x|\varrho) = G(o,o|\varrho)\prod_{k=1}^{\|x\|} L_{\tau(x_k)}\big(o_{\tau(x_k)}, x_k|\xi_i(\varrho)\big) \leq G(o,o|\varrho)\left(\frac{1}{1-\max_{i\in\mathscr{I}}\xi_i(\varrho)}\right)^{\|x\|}.$$

Since $\|X_n\| \leq n$, we obtain

$$p^{(m)}(e,X_n) \leq G(o,o|\varrho)\cdot\left(\frac{1}{1-\max_{i\in\mathscr{I}}\xi_i(\varrho)}\right)^n\cdot\varrho^{-m}.$$

□

Lemma 3.5. *Let $(A_n)_{n\in\mathbb{N}}$, $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$ be sequences of strictly positive numbers with $A_n = a_n + b_n$. Assume that $\lim_{n\to\infty} -\frac{1}{n}\log A_n = c \in [0,\infty)$ and that $\lim_{n\to\infty} b_n/q^n = 0$ for all $q \in (0,1)$. Then $\lim_{n\to\infty} -\frac{1}{n}\log a_n = c$.*

Proof. Under the made assumptions it can not be that $\liminf_{n\to\infty} a_n/q^n = 0$ for every $q \in (0,1)$. Indeed, assume that this would hold. Choose any $q > 0$. Then there is a subseqence $(a_{n_k})_{k\in\mathbb{N}}$ with $a_{n_k}/q^{n_k} \to 0$. Moreover, there is $N_q \in \mathbb{N}$ such that $a_{n_k}, b_{n_k} < q^{n_k}/2$ for all $k \geq N_q$. But this implies

$$-\frac{1}{n_k}\log(a_{n_k} + b_{n_k}) \geq -\frac{1}{n_k}\log(q^{n_k}) = -\log q.$$

The last inequality holds for every $q > 0$, yielding that $\limsup_{n\to\infty} -\frac{1}{n}\log A_n = \infty$, a contradiction. Thus, there is some $N \in \mathbb{N}$ such that $b_n < a_n$ for all $n \geq N$. We get for all $n \geq N$:

$$\begin{aligned}
-\frac{1}{n}\log(a_n + b_n) &\leq -\frac{1}{n}\log(a_n) = -\frac{1}{n}\log\Big(\frac{1}{2}a_n + \frac{1}{2}a_n\Big)\\
&\leq -\frac{1}{n}\log\Big(\frac{1}{2}a_n + \frac{1}{2}b_n\Big) \leq -\frac{1}{n}\log\frac{1}{2} - \frac{1}{n}\log(a_n + b_n).
\end{aligned}$$

Taking limits yields that $-\frac{1}{n}\log(a_n)$ tends to c, since the leftmost and rightmost side of this inequality chain tend to c. □

For the next lemma recall the definition of K from (2.1).

Lemma 3.6. *For $n \in \mathbb{N}$, consider the function $f_n : V \to \mathbb{R}$ defined by*

$$f_n(x) := \begin{cases} -\frac{1}{n}\log\sum_{m=0}^{Kn^2} p^{(m)}(o,x), & \text{if } p^{(n)}(o,x) > 0,\\ 0, & \text{otherwise.}\end{cases}$$

Then there are constants d and D such that $d \leq f_n(x) \leq D$ for all $n \in \mathbb{N}$ and $x \in V$.

Proof. Assume that $p^{(n)}(o,x) > 0$. Recall from the proof of Corollary 3.3 that we have $G(x,x|1) \leq \big(1-\frac{1}{R}\big)^{-1}$. Therefore,

$$\sum_{m=0}^{Kn^2} p^{(m)}(o,x) \leq G(o,x|1) \leq F(o,x|1)\cdot G(x,x|1) \leq \frac{1}{1-\frac{1}{R}},$$

that is

$$f_n(x) \geq -\frac{1}{n}\log\frac{1}{1-\frac{1}{R}}.$$

For the upper bound, observe that, by uniform irreducibility, $x \in V$ with $p^{(n)}(o,x) > 0$ can be reached from o in $N_x \leq K\cdot|x| \leq Kn$ steps with a probability of at least $\varepsilon_0^{|x|}$, where $\varepsilon_0 > 0$ from (2.1) is independent from x. Thus, at least one of the summands in $\sum_{m=0}^{Kn^2} p^{(m)}(o,x)$ has a value greater or equal to $\varepsilon_0^{|x|} \geq \varepsilon_0^n$. Thus, $f_n(x) \leq -\log\varepsilon_0$. □

Now we can state and prove our first main result:

Theorem 3.7. *Assume $R > 1$. Then the asymptotic entropy exists and is given by*

$$h = \ell_0 \cdot \sum_{g \in V_*^\times} l(g)\, \pi\big(g, \tau(g)\big) = \ell.$$

Proof. By (2.4) we can rewrite ℓ as

$$\ell = \lim_{n\to\infty} -\frac{1}{n} \log L(o, X_n|1) = \lim_{n\to\infty} -\frac{1}{n} \log \frac{G(o,X_n|1)}{G(o,o|1)} = \lim_{n\to\infty} -\frac{1}{n} \log G(o,X_n|1).$$

Since

$$G(o,X_n|1) = \sum_{m\geq 0} p^{(m)}(o,X_n) \geq p^{(n)}(o,X_n) = \pi_n(X_n),$$

we have

$$\liminf_{n\to\infty} -\frac{1}{n} \log \pi_n(X_n) \geq \ell. \tag{3.5}$$

The next aim is to prove $\limsup_{n\to\infty} -\frac{1}{n}\mathbb{E}\big[\log \pi_n(X_n)\big] \leq \ell$. We now apply Lemma 3.5 by setting

$$A_n := \sum_{m\geq 0} p^{(m)}(o,X_n),\ a_n := \sum_{m=0}^{Kn^2} p^{(m)}(o,X_n) \text{ and } b_n := \sum_{m\geq Kn^2+1} p^{(m)}(o,X_n).$$

By Lemma 3.4,

$$b_n \leq \sum_{m\geq Kn^2+1} \frac{G(o,o|\varrho)}{\varrho^m} \cdot \Big(\frac{1}{1-\max_{i\in\mathscr{I}} \xi_i(\varrho)}\Big)^n = G(o,o|\varrho)\cdot \Big(\frac{1}{1-\max_{i\in\mathscr{I}} \xi_i(\varrho)}\Big)^n \cdot \frac{\varrho^{-Kn^2-1}}{1-\varrho^{-1}}.$$

Therefore, b_n decays faster than any geometric sequence. Applying Lemma 3.5 yields

$$\ell = \lim_{n\to\infty} -\frac{1}{n} \log \sum_{m=0}^{Kn^2} p^{(m)}(o,X_n) \quad \text{almost surely.}$$

By Lemma 3.6, we may apply the Dominated Convergence Theorem and get:

$$\begin{aligned}
\ell &= \int \lim_{n\to\infty} -\frac{1}{n} \log \sum_{m=0}^{Kn^2} p^{(m)}(o,X_n)\, d\mathbb{P} \\
&= \lim_{n\to\infty} \int -\frac{1}{n} \log \sum_{m=0}^{Kn^2} p^{(m)}(o,X_n)\, d\mathbb{P} \\
&= \lim_{n\to\infty} -\frac{1}{n} \sum_{x\in V} p^{(n)}(o,x) \log \sum_{m=0}^{Kn^2} p^{(m)}(o,x).
\end{aligned}$$

Recall that *Shannon's Inequality* gives

$$-\sum_{x\in V} p^{(n)}(o,x) \log \mu(x) \geq -\sum_{x\in V} p^{(n)}(o,x) \log p^{(n)}(o,x)$$

for every finitely supported probability measure μ on V. We apply now this inequality by setting $\mu(x) := \frac{1}{Kn^2+1}\sum_{m=0}^{Kn^2} p^{(m)}(o,x)$:

$$\begin{aligned}\ell &\geq \limsup_{n\to\infty} \frac{1}{n}\sum_{x\in V} p^{(n)}(o,x)\log(Kn^2+1) - \frac{1}{n}\sum_{x\in V} p^{(n)}(o,x)\log p^{(n)}(o,x)\\ &= \limsup_{n\to\infty} -\frac{1}{n}\int \log \pi_n(X_n)\,d\mathbb{P}.\end{aligned}$$

Now we can conclude with Fatou's Lemma:

$$\begin{aligned}\ell \leq \int \liminf_{n\to\infty} \frac{-\log \pi_n(X_n)}{n}\,d\mathbb{P} &\leq \liminf_{n\to\infty} \int \frac{-\log \pi_n(X_n)}{n}\,d\mathbb{P}\\ &\leq \limsup_{n\to\infty} \int \frac{-\log \pi_n(X_n)}{n}\,d\mathbb{P} \leq \ell. \end{aligned} \tag{3.6}$$

Thus, $\lim_{n\to\infty} -\frac{1}{n}\mathbb{E}\big[\log \pi_n(X_n)\big]$ exists and the limit equals ℓ. The rest follows from (3.2) and (3.3). □

We now give another formula for the asymptotic entropy which shows that it is strictly positive.

Theorem 3.8. *Assume $R > 1$. Then the asymptotic entropy is given by*

$$h = \ell_0 \cdot \sum_{g,h\in V_*^\times} -\pi\big(g,\tau(g)\big)\, q\big((g,\tau(g)),(h,\tau(h))\big)\, \log q\big((g,\tau(g)),(h,\tau(h))\big) > 0.$$

Remarks: Observe that the sum on the right hand side of Theorem 3.8 equals the entropy rate (for positive recurrent Markov chains) of $\big(\widetilde{\mathbf{W}}_k, \tau_k\big)_{k\in\mathbb{N}}$, which is defined by the almost sure constant limit

$$h_Q := \lim_{n\to\infty} -\frac{1}{n}\log \mu_n\big((\widetilde{\mathbf{W}}_1,\tau_1),\dots,(\widetilde{\mathbf{W}}_n,\tau_n)\big),$$

where $\mu_n\big((g_1,\tau_1),\dots,(g_n,\tau_n)\big)$ is the joint distribution of $\big((\widetilde{\mathbf{W}}_1,\tau_1),\dots,(\widetilde{\mathbf{W}}_n,\tau_n)\big)$. That is, $h = \ell\cdot h_Q$. For more details, we refer e.g. to Cover and Thomas [6, Chapter 4].

At this point it is essential that we have defined the length function $l(\cdot)$ with the help of the functions $L(x,y|z)$ and not by the Greenian metric.

Proof. For a moment let be $x = x_1\dots x_n \in V$. Then:

$$\begin{aligned} l(x) &= -\log \prod_{j=1}^{n} L_{\tau(x_j)}\big(o_{\tau(x_j)}, x_j | \xi_{\tau(x_j)}\big)\\ &= -\log \prod_{j=2}^{n} \frac{\alpha_{\tau(x_j)}}{\alpha_{\tau(x_{j-1})}} \frac{\xi_{\tau(x_{j-1})}}{\xi_{\tau(x_j)}} \frac{1-\xi_{\tau(x_j)}}{1-\xi_{\tau(x_{j-1})}} L_{\tau(x_j)}\big(o_{\tau(x_j)}, x_j|\xi_{\tau(x_j)}\big)\\ &\qquad -\log L_{\tau(x_1)}\big(o_{\tau(x_1)}, x_1|\xi_{\tau(x_1)}\big) + \log \frac{\xi_{\tau(x_1)}\,\alpha_{\tau(x_n)}\,(1-\xi_{\tau(x_n)})}{\alpha_{\tau(x_1)}\,\xi_{\tau(x_n)}\,(1-\xi_{\tau(x_1)})}\\ &= -\log \prod_{j=2}^{n} q\big((x_{j-1},\tau(x_{j-1})),(x_j,\tau(x_j))\big)\\ &\qquad -\log L_{\tau(x_1)}\big(o_{\tau(x_1)}, x_1|\xi_{\tau(x_1)}\big) + \log \frac{\xi_{\tau(x_1)}\,\alpha_{\tau(x_n)}\,(1-\xi_{\tau(x_n)})}{\alpha_{\tau(x_1)}\,\xi_{\tau(x_n)}\,(1-\xi_{\tau(x_1)})}. \end{aligned} \tag{3.7}$$

We now replace x by $X_{\mathbf{e}_k}$ in the last equation: since $l(X_n)/n$ tends to h almost surely, the subsequence $\bigl(l(X_{\mathbf{e}_k})/\mathbf{e}_k\bigr)_{k\in\mathbb{N}}$ converges also to h. Since $\min_{i\in\mathcal{I}}\xi_i > 0$ and $\max_{i\in\mathcal{I}}\xi_i < 1$, we get

$$\frac{1}{\mathbf{e}_k}\log\frac{\xi_{\tau(x_1)}\,\alpha_{\tau(x_k)}\,(1-\xi_{\tau(x_k)})}{\alpha_{\tau(x_1)}\,\xi_{\tau(x_k)}\,(1-\xi_{\tau(x_1)})}\xrightarrow{k\to\infty} 0 \quad \text{almost surely,}$$

where $x_1 := X_{\mathbf{e}_1}$ and $x_k := \widetilde{\mathbf{W}}_k = \widetilde{X}_{\mathbf{e}_k}$. By positive recurrence of $\bigl(\widetilde{\mathbf{W}}_k,\tau_k\bigr)_{k\in\mathbb{N}}$, an application of the ergodic theorem yields

$$\begin{aligned}
&-\frac{1}{k}\log\prod_{j=2}^{k} q\bigl((\widetilde{\mathbf{W}}_{j-1},\tau_{j-1}),(\widetilde{\mathbf{W}}_j,\tau_j)\bigr)\\
\xrightarrow{n\to\infty}\ & h' := -\sum_{\substack{g,h\in V_*^\times;\\ \tau(g)\neq\tau(h)}} \pi\bigl(g,\tau(g)\bigr)\, q\bigl((g,\tau(g)),(h,\tau(h)\bigr)\log q\bigl((g,\tau(g)),(h,\tau(h)\bigr) > 0 \text{ a.s.,}
\end{aligned}$$

whenever $h' < \infty$. Obviously, for every $x_1 \in V_*^\times$

$$\lim_{k\to\infty} -\frac{1}{\mathbf{e}_k}\log L_{\tau(x_1)}\bigl(o_{\tau(x_1)},x_1|\xi_{\tau(x_1)}\bigr) = 0 \quad \text{almost surely.}$$

Since $\lim_{k\to\infty} k/\mathbf{e}_k = \ell_0$ we get

$$h = \lim_{k\to\infty}\frac{l\bigl(X_{\mathbf{e}_k}\bigr)}{\mathbf{e}_k} = \lim_{k\to\infty}\frac{l\bigl(X_{\mathbf{e}_k}\bigr)}{k}\frac{k}{\mathbf{e}_k} = h'\cdot\ell_0,$$

whenever $h' < \infty$. In particular, $h > 0$ since $\ell_0 > 0$ by [11, Section 4].

It remains to show that it cannot be that $h' = \infty$. For this purpose, assume now $h' = \infty$. Define for $N \in \mathbb{N}$ the function $h_N : \bigl(V_*^\times\bigr)^2 \to \mathbb{R}$ by

$$h_N(g,h) := N \wedge \bigl(-\log q\bigl((g,\tau(g)),(h,\tau(h))\bigr).$$

Then

$$\begin{aligned}
&-\frac{1}{k}\sum_{j=2}^{k}\log h_N\bigl(\widetilde{X}_{\mathbf{e}_{j-1}},\widetilde{X}_{\mathbf{e}_j}\bigr)\\
\xrightarrow{k\to\infty}\ & h'_N := -\sum_{\substack{g,h\in V_*^\times,\\ \tau(g)\neq\tau(h)}} \pi\bigl(g,\tau(g)\bigr)\, q\bigl((g,\tau(g)),(h,\tau(h))\bigr)\log h_N(g,h) \quad \text{almost surely.}
\end{aligned}$$

Observe that $h'_N \to \infty$ as $N \to \infty$. Since $h_N(g,h) \leq -\log q\bigl((g,\tau(g)),(h,\tau(h))\bigr)$ and $h' = \infty$ by assumption there is for every $M \in \mathbb{R}$ and almost every trajectory of $\bigl(\widetilde{\mathbf{W}}_k\bigr)_{k\in\mathbb{N}}$ an almost surely finite random time $\mathbf{T_q} \in \mathbb{N}$ such that for all $k \geq \mathbf{T_q}$

$$-\frac{1}{k}\sum_{j=2}^{k}\log q\bigl((\widetilde{\mathbf{W}}_{j-1},\tau_{j-1}),(\widetilde{\mathbf{W}}_j,\tau_j)\bigr) > M. \tag{3.8}$$

On the other hand side there is for every $M > 0$, every small $\varepsilon > 0$ and almost every trajectory an almost surely finite random time $\mathbf{T_L}$ such that for all $k \geq \mathbf{T_L}$

$$
\begin{aligned}
& -\frac{1}{\mathbf{e}_k}\sum_{j=1}^{k}\log L_{\tau(X_{\mathbf{e}_j})}\big(o_{\tau(X_{\mathbf{e}_j})},\widetilde{X}_{\mathbf{e}_j}|\xi_{\tau(X_{\mathbf{e}_j})}\big) \in (h-\varepsilon,h+\varepsilon) \quad \text{and} \\
& -\frac{1}{\mathbf{e}_k}\sum_{j=2}^{k}\log q\big((\widetilde{X}_{\mathbf{e}_{j-1}},\tau_{j-1}),(\widetilde{X}_{\mathbf{e}_j},\tau_j)\big) \\
= \; & -\frac{k}{\mathbf{e}_k}\frac{1}{k}\sum_{j=2}^{k}\log q\big((\widetilde{X}_{\mathbf{e}_{j-1}},\tau_{j-1}),(\widetilde{X}_{\mathbf{e}_j},\tau_j)\big) > \ell_0\cdot M.
\end{aligned}
$$

Furthermore, since $\min_{i\in\mathcal{I}}\xi_i > 0$ and $\max_{i\in\mathcal{I}}\xi_i < 1$ there is an almost surely finite random time $\mathbf{T}_\varepsilon \geq \mathbf{T_L}$ such that for all $k \geq \mathbf{T}_\varepsilon$ and all $x_1 = X_{\mathbf{e}_1}$ and $x_k = \widetilde{X}_{\mathbf{e}_k}$

$$
\begin{aligned}
& -\frac{1}{\mathbf{e}_k}\log\frac{\xi_{\tau(x_1)}\,\alpha_{\tau(x_k)}\,(1-\xi_{\tau(x_k)})}{\alpha_{\tau(x_1)}\,\xi_{\tau(x_k)}\,(1-\xi_{\tau(x_1)})} \in (-\varepsilon,\varepsilon) \quad \text{and} \\
& \frac{1}{\mathbf{e}_k}\log L_{\tau(x_1)}\big(o_{\tau(x_1)},x_1|\xi_{\tau(x_1)}\big) \in (-\varepsilon,\varepsilon).
\end{aligned}
$$

Choose now $M > (h+3\varepsilon)/\ell_0$. Then we get the desired contradiction, when we substitute in equality (3.7) the vertex x by $X_{\mathbf{e}_k}$ with $k \geq \mathbf{T}_\varepsilon$, divide by $\mathbf{e}_k$ on both sides and see that the left side is in $(h-\varepsilon,h+\varepsilon)$ and the rightmost side is bigger than $h+\varepsilon$. This finishes the proof of Theorem 3.8. □

Corollary 3.9. *Assume $R > 1$. Then we have for almost every path of the random walk $(X_n)_{n\in\mathbb{N}_0}$*

$$h = \liminf_{n\to\infty} -\frac{\log \pi_n(X_n)}{n}.$$

Proof. Recall Inequality (3.5). Integrating both sides of this inequality yields together with the inequality chain (3.6) that

$$\int \liminf_{n\to\infty} -\frac{\log \pi_n(X_n)}{n} - h\, d\mathbb{P} = 0,$$

providing that $h = \liminf_{n\to\infty} -\frac{1}{n}\log \pi_n(X_n)$ for almost every realisation of the random walk. □

The following lemma gives some properties concerning general measure theory:

Lemma 3.10. *Let $(Z_n)_{n\in\mathbb{N}_0}$ be a sequence of non-negative random variables and $0 < c \in \mathbb{R}$. Suppose that $\liminf_{n\to\infty} Z_n \geq c$ almost surely and $\lim_{n\to\infty}\mathbb{E}[Z_n] = c$. Then the following holds:*

1. *$Z_n \xrightarrow{\mathbb{P}} c$, that is, Z_n converges in probability to c.*
2. *If Z_n is uniformly bounded then $Z_n \xrightarrow{L_1} c$, that is, $\int \big|Z_n - c\big| d\mathbb{P} \to 0$ as $n \to \infty$*

Proof. First, we prove convergence in probability of $(Z_n)_{n\in\mathbb{N}_0}$. For every $\delta_1 > 0$, there is some index N_{δ_1} such that for all $n \geq N_{\delta_1}$

$$\int Z_n \, d\mathbb{P} \in (c-\delta_1, c+\delta_1).$$

Furthermore, due to the above made assumptions on $(Z_n)_{n\in\mathbb{N}_0}$ there is for every $\delta_2 > 0$ some index N_{δ_2} such that for all $n \geq N_{\delta_2}$

$$\mathbb{P}[Z_n > c-\delta_1] > 1-\delta_2. \tag{3.9}$$

Since $c = \lim_{n\to\infty} \int Z_n \, d\mathbb{P}$ it must be that for every arbitrary but fixed $\varepsilon > 0$, every $\delta_1 < \varepsilon$ and for all n big enough

$$\mathbb{P}\big[\, Z_n > c-\delta_1\big] \cdot (c-\delta_1) + \mathbb{P}\big[\, Z_n > c+\varepsilon\big] \cdot (\varepsilon+\delta_1) \leq \int Z_n \, d\mathbb{P} \leq c+\delta_1,$$

or equivalently,

$$\mathbb{P}\big[\, Z_n > c+\varepsilon\big] \leq \frac{c+\delta_1 - \mathbb{P}\big[\, Z_n > c-\delta_1\big]\cdot(c-\delta_1)}{\varepsilon+\delta_1}.$$

Letting $\delta_2 \to 0$ we get

$$\limsup_{n\to\infty} \mathbb{P}\big[\, Z_n > c+\varepsilon\big] \leq \frac{2\delta_1}{\varepsilon+\delta_1}.$$

Since we can choose δ_1 arbitrarily small we get

$$\mathbb{P}\big[\, Z_n > c+\varepsilon\big] \xrightarrow{n\to\infty} 0 \quad \text{for all } \varepsilon > 0.$$

This yields convergence in probability of Z_n to c.

In order to prove the second part of the lemma we define for any small $\varepsilon > 0$ and $n \in \mathbb{N}$ the events

$$A_{n,\varepsilon} := \big[\, |Z_n - c| \leq \varepsilon\big] \text{ and } B_{n,\varepsilon} := \big[\, |Z_n - c| > \varepsilon\big].$$

For arbitrary but fixed $\varepsilon > 0$, convergence in probability of Z_n to c gives an integer $N_\varepsilon \in \mathbb{N}$ such that $\mathbb{P}[B_{n,\varepsilon}] < \varepsilon$ for all $n \geq N_\varepsilon$. Since $0 \leq Z_n \leq M$ is assumed to be uniformly bounded, we get for $n \geq N_\varepsilon$:

$$\int |Z_n - c| \, d\mathbb{P} = \int_{A_{n,\varepsilon}} |Z_n - c| \, d\mathbb{P} + \int_{B_{n,\varepsilon}} |Z_n - c| \, d\mathbb{P} \leq \varepsilon + \varepsilon(M+c) \xrightarrow{\varepsilon\to 0} 0.$$

Thus, we have proved the second part of the lemma. □

We can apply the last lemma immediately to our setting:

Corollary 3.11. *Assume $R > 1$. Then we have the following types of convergence:*

1. *Convergence in probability:*

$$-\frac{1}{n}\log \pi_n(X_n) \xrightarrow{\mathbb{P}} h.$$

2. *Assume that there is $c_0 > 0$ such that $p(x,y) \geq c_0$ whenever $p(x,y) > 0$. Then:*

$$-\frac{1}{n}\log \pi_n(X_n) \xrightarrow{L_1} h.$$

Proof. Setting $Z_n = -\frac{1}{n}\log \pi_n(X_n)$ and applying Lemma 3.10 yields the claim. Note that the assumption $p(x,y) \geq c_0$ yields $0 \leq \frac{-\log \pi_n(X_n)}{n} \leq -\log c_0$. □

The assumption of the second part of the last corollary is obviously satisfied if we consider free products of *finite* graphs.

The reasoning in our proofs for existence of the entropy and its different properties (in particular, the reasoning in Section 3.2) is very similar to the argumentation in [12]. However, the structure of free products of graphs is more complicated than in the case of directed covers as considered in [12]. We outline the main differences to the reasoning in the aforementionend article. First, in [12] a very similar rate of escape (compare [12, Theorem 3.8] with Proposition 3.2) is considered, which arises from a length function induced by last visit generating functions. While the proof of existence of the rate of escape in [12] is easy to check, we have to make more effort in the case of free products, since $-\log L_i(o_i, x|1)$ is not necessarily bounded for $x \in V_i$. Additionally, one has to study the various ingridients of the proof more carefully, since non-trivial loops are possible in our setting in contrast to random walks on trees. Secondly, in [12] the invariant measure $\pi\big(g, \tau(g)\big)$ of our proof collapses to $\nu\big(\tau(g)\big)$, that is, in [12] one has to study the sequence $\big(\tau(\mathbf{W}_k)\big)_{k\in\mathbb{N}}$, while in our setting we have to study the more complex sequence $\big(\widetilde{\mathbf{W}}_k, \tau(\mathbf{W}_k)\big)_{k\in\mathbb{N}}$; compare [12, proof of Theorem 3.8] with Lemma 3.1 and Proposition 3.2.

4 A Formula via Double Generating Functions

In this section we derive another formula for the asymptotic entropy. The main tool is the following theorem of Sawyer and Steger [25, Theorem 2.2]:

Theorem 4.1 (Sawyer and Steger). *Let $(Y_n)_{n\in\mathbb{N}_0}$ be a sequence of real-valued random variables such that, for some $\delta > 0$,*

$$\mathbb{E}\left(\sum_{n\geq 0} \exp(-rY_n - sn)\right) = \frac{C(r,s)}{g(r,s)} \quad \text{for } 0 < r,s < \delta,$$

where $C(r,s)$ and $g(r,s)$ are analytic for $|r|,|s| < \delta$ and $C(0,0) \neq 0$. Denote by g'_r and g'_s the partial derivatives of g with respect to r and s. Then

$$\frac{Y_n}{n} \xrightarrow{n\to\infty} \frac{g'_r(0,0)}{g'_s(0,0)} \quad \text{almost surely.}$$

Setting $z = e^{-s}$ and $Y_n := -\log L(o, X_n|1)$ the expectation in Theorem 4.1 becomes

$$\mathscr{E}(r,z) = \sum_{x\in V}\sum_{n\geq 0} p^{(n)}(o,x) L(o,x|1)^r z^n = \sum_{x\in V} G(o,x|z) L(o,x|1)^r.$$

We define for $i \in \mathscr{I}$, $r,z \in \mathbb{C}$:

$$\begin{aligned}
\mathscr{L}(r,z) &:= 1 + \sum_{n\geq 1} \sum_{x_1 \ldots x_n \in V\setminus\{o\}} \prod_{j=1}^{n} L_{\tau(x_j)}\big(o_{\tau(x_j)}, x_j | \xi_{\tau(x_j)}(z)\big) \cdot L_{\tau(x_j)}\big(o_{\tau(x_j)}, x_j | \xi_{\tau(x_j)}\big)^r, \\
\mathscr{L}_i^+(r,z) &:= \sum_{x\in V_i^\times} L_i\big(o_i, x|\xi_i(z)\big) L_i(o_i, x|\xi_i)^r.
\end{aligned}$$

Finally, $\mathcal{L}_i(r,z)$ is defined by

$$\mathcal{L}_i^+(r,z)\cdot\left(1+\sum_{n\geq 2}\sum_{\substack{x_2\ldots x_n\in V^\times\setminus\{o\},\\ \tau(x_2)\neq i}}\prod_{j=2}^{n} L_{\tau(x_j)}\big(o_{\tau(x_j)},x_j|\xi_{\tau(x_j)}(z)\big)\cdot L_{\tau(x_j)}\big(o_{\tau(x_j)},x_j|\xi_{\tau(x_j)}\big)^r\right).$$

With these definitions we have $\mathcal{L}(r,z)=1+\sum_{i\in\mathcal{I}}\mathcal{L}_i(r,z)$ and $\mathcal{E}(r,z)=G(o,o|z)\cdot\mathcal{L}(r,z)$. Simple computations analogously to [11, Lemma 4.2, Corollary 4.3] yield

$$\mathcal{E}(r,z)=\frac{G(o,o|z)}{1-\mathcal{L}^*(r,z)},\ \text{where } \mathcal{L}^*(r,z)=\sum_{i\in\mathcal{I}}\frac{\mathcal{L}_i^+(r,z)}{1+\mathcal{L}_i^+(r,z)}.$$

We now define $C(r,z):=G(o,o|z)$ and $g(r,z):=1-\mathcal{L}^*(r,z)$ and apply Theorem 4.1 by differentiating $g(r,z)$ and evaluating the derivatives at $(0,1)$:

$$\begin{aligned}
\frac{\partial g(r,z)}{\partial r}\bigg|_{r=0,z=1} &= -\sum_{i\in\mathcal{I}}\frac{\sum_{x\in V_i^\times}L_i(o_i,x|\xi_i)\cdot\log L_i(o_i,x|\xi_i)}{\big(1+\sum_{x\in V_i^\times}L_i(o_i,x|\xi_i)\big)^2}\\
&= -\sum_{i\in\mathcal{I}}G_i(o_i,o_i|\xi_i)\cdot(1-\xi_i)^2\cdot\sum_{x\in V_i^\times}G_i(o_i,x|\xi_i)\log L_i(o_i,x|\xi_i)\\
&= -\sum_{i\in\mathcal{I}}G_i(o_i,o_i|\xi_i)\cdot(1-\xi_i)^2\cdot\\
&\qquad\cdot\left(\sum_{x\in V_i}G_i(o_i,x|\xi_i)\log G_i(o_i,x|\xi_i)-\frac{\log G_i(o_i,o_i|\xi_i)}{1-\xi_i}\right),\\
\frac{\partial g(r,z)}{\partial s}\bigg|_{r=0,s=0} &= \sum_{i\in\mathcal{I}}\frac{\partial}{\partial z}\Big(1-\big(1-\xi_i(z)\big)G_i\big(o_i,o_i|\xi_i(z)\big)\Big)\Big|_{z=1}\\
&= \sum_{i\in\mathcal{I}}\xi_i'(1)\cdot\big(G_i(o_i,o_i|\xi_i)-(1-\xi_i)G_i'(o_i,o_i|\xi_i)\big).
\end{aligned}$$

Corollary 4.2. *Assume $R>1$. Then the entropy can be rewritten as*

$$h=\frac{\frac{\partial g(r,z)}{\partial r}(0,1)}{\frac{\partial g(r,z)}{\partial s}(0,1)}.$$

□

5 Entropy of Random Walks on Free Products of Groups

In this section let each V_i be a finitely generated group Γ_i with identity $e_i=o_i$. W.l.o.g. we assume that the V_i's are pairwise disjoint. The free product is again a group with concatenation (followed by iterated cancellations and contractions) as group operation. We write $\Gamma_i^\times:=\Gamma_i\setminus\{e_i\}$. Suppose we are given a probability measure μ_i on $\Gamma_i\setminus\{e_i\}$ for every $i\in\mathcal{I}$ governing a random walk on Γ_i, that is, $p_i(x,y)=\mu_i(x^{-1}y)$ for all $x,y\in\Gamma_i$. Let $(\alpha_i)_{i\in\mathcal{I}}$ be a family of strictly positive real numbers with $\sum_{i\in\mathscr{I}}\alpha_i = 1$. Then the random walk on the free product $\Gamma := \Gamma_1 * \cdots * \Gamma_r$ is defined by the transition probabilities $p(x,y) = \mu(x^{-1}y)$, where

$$\mu(w) = \begin{cases} \alpha_i\mu_i(w), & \text{if } w \in \Gamma_i^\times, \\ 0, & \text{otherwise.} \end{cases}$$

Analogously, $\mu^{(n)}$ denotes the n-th convolution power of μ. The random walk on Γ starting at the identity e of Γ is again denoted by the sequence $(X_n)_{n\in\mathbb{N}_0}$. In particular, the radius of convergence of the associated Green function is strictly bigger than 1; see [29, Theorem 10.10, Corollary 12.5]. In the case of free products of groups it is well-known that the entropy exists and can be written as

$$h = \lim_{n\to\infty}\frac{-\log\pi_n(X_n)}{n} = \lim_{n\to\infty}\frac{-\log F(e,X_n|1)}{n};$$

see Derriennic [7], Kaimanovich and Vershik [14] and Blachère, Haïssinsky and Mathieu [3]. For free products of *finite* groups, Mairesse and Mathéus [21] give an explicit formula for h, which remains also valid for free products of countable groups, but in the latter case one needs the solution of an infinite system of polynomial equations. In the following we will derive another formula for the entropy, which holds also for free products of *infinite* groups.

We set $l(g_1\dots g_n) := -\log F(e,g_1\dots g_n|1)$. Observe that transitivity yields $F(g,gh|1) = F(e,h|1)$. Thus,

$$l(g_1\dots g_n) = -\log\prod_{j=1}^n F(g_1\dots g_{j-1}, g_1\dots g_j|1) = -\sum_{j=1}^n \log F(e,g_j|1).$$

First, we rewrite the following expectations as

$$\begin{aligned}\mathbb{E}l(X_n) &= \sum_{i\in\mathscr{I}}\sum_{g\in\Gamma_i}\alpha_i\mu_i(g)\sum_{h\in\Gamma} l(h)\,\mu^{(n)}(h),\\ \mathbb{E}l(X_{n+1}) &= \sum_{i\in\mathscr{I}}\sum_{g\in\Gamma_i}\alpha_i\mu_i(g)\sum_{h\in\Gamma} l(gh)\,\mu^{(n)}(h).\end{aligned}$$

Thus,

$$\begin{aligned}\mathbb{E}l(X_{n+1}) - \mathbb{E}l(X_n) &= \sum_{i\in\mathscr{I}}\sum_{g\in\Gamma_i}\alpha_i\mu_i(g)\int\big(l(gh)-l(h)\big)\,d\mu^{(n)}(h)\\ &= \sum_{i\in\mathscr{I}}\sum_{g\in\Gamma_i}\alpha_i\mu_i(g)\int -\log\frac{F(e,gX_n|1)}{F(e,X_n|1)}\,d\mu^{(n)}. \qquad (5.1)\end{aligned}$$

Recall that $\|X_n\| \to \infty$ almost surely. That is, X_n converges almost surely to a random infinite word X_∞ of the form $x_1x_2\ldots \in \Big(\bigcup_{i=1}^r \Gamma_i^\times\Big)^{\mathbb{N}}$, where two consecutive letters are not from the same $\Gamma_i^\times$. Denote by $X_\infty^{(1)}$ the first letter of X_∞. Let be $g \in \Gamma_i^\times$. For $n \geq \mathbf{e}_1$, the integrand in (5.1) is constant: if $\tau\big(X_\infty^{(1)}\big) \neq i$ then

$$\log\frac{F(e,gX_n|1)}{F(e,X_n|1)} = \log F(e,g),$$

and if $\tau\big(X_\infty^{(1)}\big) = i$ then

$$\log\frac{F(e,gX_n|1)}{F(e,X_n|1)} = \log\frac{F\big(e,gX_\infty^{(1)}|1\big)}{F\big(e,X_\infty^{(1)}|1\big)}.$$

By [11, Section 5], for $i \in \mathcal{I}$ and $g \in \Gamma_i^\times$,

$$\begin{aligned} \varrho(i) &:= \mathbb{P}[X_\infty^{(1)} \in \Gamma_i] = 1-(1-\xi_i)\, G_i(o_i,o_i|\xi_i) \quad \text{and} \\ \mathbb{P}[X_\infty^{(1)} = g] &= F(o_i,g|\xi_i)(1-\xi_i)\, G_i(o_i,o_i|\xi_i) = (1-\xi_i)\, G_i(o_i,g|\xi_i). \end{aligned}$$

Recall that $F(e,g|1) = F_i(o_i,g|\xi_i)$ for each $g \in \Gamma_i$. We get:

Theorem 5.1. *Whenever $h_i := -\sum_{g\in\Gamma_i} \mu_i(g)\log\mu_i(g) < \infty$ for all $i \in \mathcal{I}$, that is, when all random walks on the factors Γ_i have finite single-step entropy, then the asymptotic entropy h of the random walk on Γ is given by*

$$h = -\sum_{i\in\mathcal{I}}\sum_{g\in\Gamma_i} \alpha_i\mu_i(g)\Big[\big(1-\varrho(i)\big)\,\log F_i(o_i,g|\xi_i) + (1-\xi_i)\, G_i(o_i,o_i|\xi_i)\,\mathcal{F}(g)\Big],$$

where

$$\mathcal{F}(g) := \sum_{g'\in\Gamma_i^\times} F_i(o_i,g'|\xi_i)\log\frac{F_i(o_i,gg'|\xi_i)}{F_i(o_i,g'|\xi_i)} \text{ for } g\in\Gamma_i. \tag{5.2}$$

Proof. Consider the sequence $\mathbb{E}l(X_{n+1}) - \mathbb{E}l(X_n)$. If this sequence converges, its limit must equal h. By the above made considerations we get

$$\begin{aligned} &\mathbb{E}l(X_{n+1}) - \mathbb{E}l(X_n) \\ \xrightarrow{n\to\infty}\ & -\sum_{i\in\mathcal{I}}\sum_{g\in\Gamma_i}\mu(g)\left[(1-\varrho(i))\log F_i(o_i,g|\xi_i) + \sum_{g'\in\Gamma_i^\times}\mathbb{P}[X_\infty^{(1)} = g']\log\frac{F_i(o_i,gg'|\xi_i)}{F_i(o_i,g'|\xi_i)}\right], \end{aligned}$$

if the sum on the right hand side is finite. We have now established the proposed formula, but it remains to verify finiteness of the sum above. This follows from the following observations:
Claim A: $-\sum_{i\in\mathcal{I}}\sum_{g\in\Gamma_i}\alpha_i\mu_i(g)(1-\varrho(i))\log F_i(o_i,g|\xi_i)$ is finite.
Observe that $F_i(o_i,g|\xi_i) \geq \mu_i(g)\xi_i$ for $g \in \mathrm{supp}(\mu_i)$. Thus,

$$0 < -\sum_{g\in\Gamma_i}\mu_i(g)\log F_i(o_i,g|\xi_i) \leq -\sum_{g\in\Gamma_i}\mu_i(g)\log\big(\mu_i(g)\,\xi_i\big) = h_i - \log\xi_i.$$

This proves Claim A.

Claim B: $\sum_{i\in\mathcal{I}}\sum_{g\in\Gamma_i}\alpha_i\mu_i(g)(1-\xi_i)\sum_{g'\in\Gamma_i^\times} G_i(o_i,g'|\xi_i)\left|\log\frac{F_i(o_i,gg'|\xi_i)}{F_i(o_i,g'|\xi_i)}\right|$ is finite.
Observe that $F_i(o_i,gg'|\xi_i)/F_i(o_i,g'|\xi_i) = G_i(o_i,gg'|\xi_i)/G_i(o_i,g'|\xi_i)$. Obviously,

$$\mu_i^{(n)}(g')\,\xi_i^n \leq G_i(o_i,g'|\xi_i) \leq \frac{1}{1-\xi_i} \quad \text{for every } n\in\mathbb{N} \text{ and } g'\in\Gamma_i.$$

For $g \in \Gamma$ set $N(g) := \{n \in \mathbb{N}_0 \mid \mu^{(n)}(g) > 0\}$. Then:

$$\begin{aligned}
0 &< \sum_{g' \in \Gamma_i^\times} \mathbb{P}[X_\infty^{(1)} = g'] \cdot \bigl|\log G_i(o_i, g'|\xi_i)\bigr| \\
&= \sum_{g' \in \Gamma_i^\times} (1-\xi_i) \cdot G_i(o_i, g'|\xi_i) \cdot \bigl|\log G_i(o_i, g'|\xi_i)\bigr| \\
&= \sum_{g' \in \Gamma_i^\times} (1-\xi_i) \cdot \sum_{n \in N(g')} \mu_i^{(n)}(g') \cdot \xi_i^n \cdot \bigl|\log G_i(o_i, g'|\xi_i)\bigr| \\
&\leq \sum_{g' \in \Gamma_i^\times} (1-\xi_i) \cdot \sum_{n \in N(g')} \mu_i^{(n)}(g') \cdot \xi_i^n \cdot \max\bigl\{-\log\bigl(\mu_i^{(n)}(g')\,\xi_i^n\bigr), -\log(1-\xi_i)\bigr\} \\
&\leq (1-\xi_i) \cdot \sum_{n \in N(g')} n\,\xi_i^n \cdot \underbrace{\frac{-1}{n} \sum_{g' \in \Gamma_i} \mu_i^{(n)}(g') \log \mu_i^{(n)}(g')}_{(*)} - (1-\xi_i)\log \xi_i \sum_{n \geq 1} n\,\xi_i^n \\
&\quad -(1-\xi_i)\log(1-\xi_i) \sum_{n \geq 1} \xi_i^n.
\end{aligned}$$

Recall that $h_i < \infty$ together with Kingman's subadditive ergodic theorem implies existence of a constant $H_i \geq 0$ with

$$\lim_{n\to\infty} -\frac{1}{n} \sum_{g \in \Gamma_i} \mu_i^{(n)}(g) \log \mu_i^{(n)}(g) = H_i. \tag{5.3}$$

Thus, if $n \in \mathbb{N}$ is large enough, the sum $(*)$ is in the interval $(H_i - \varepsilon, H_i + \varepsilon)$ for any arbitrarily small $\varepsilon > 0$. That is, the sum $(*)$ is uniformly bounded for all $n \in \mathbb{N}$. From this follows that the rightmost side of the last inequality chain is finite.

Furthermore, we have for each $g \in \Gamma_i$ with $\mu_i^{(n)}(g) > 0$:

$$\begin{aligned}
0 &< \sum_{g' \in \Gamma_i^\times} \mathbb{P}[X_\infty^{(1)} = g'] \cdot \bigl|\log G_i(o_i, gg'|\xi_i)\bigr| \\
&= \sum_{g' \in \Gamma_i^\times} (1-\xi_i) \cdot G_i(o_i, g'|\xi_i) \cdot \bigl|\log G_i(o_i, gg'|\xi_i)\bigr| \\
&= \sum_{g' \in \Gamma_i^\times} (1-\xi_i) \cdot \sum_{n \in N(g')} \mu_i^{(n)}(g') \cdot \xi_i^n \cdot \bigl|\log G_i(o_i, gg'|\xi_i)\bigr| \\
&\leq \sum_{g' \in \Gamma_i^\times} (1-\xi_i) \cdot \sum_{n \in N(g')} \mu_i^{(n)}(g') \cdot \xi_i^n \cdot \max\bigl\{-\log\bigl(\mu_i(g)\,\mu_i^{(n)}(g')\,\xi_i^{n+1}\bigr), -\log(1-\xi_i)\bigr\} \\
&\leq -(1-\xi_i) \cdot \sum_{n \in N(g')} \xi_i^n \cdot \sum_{g' \in \Gamma_i} \mu_i^{(n)}(g') \log \mu_i^{(n)}(g') - (1-\xi_i) \cdot \log \xi_i \cdot \sum_{n \geq 1} (n+1)\,\xi_i^n \\
&\quad - \log \mu_i(g) - \log(1-\xi_i).
\end{aligned}$$

If we sum up over all g with $\mu(g) > 0$, we get:

$$\underbrace{-\sum_{i\in\mathscr{I}}\sum_{g\in\Gamma_i}\alpha_i\mu_i(g)(1-\xi_i)\sum_{n\in N(g')}\xi_i^n\sum_{g'\in\Gamma_i}\mu_i^{(n)}(g')\log\mu_i^{(n)}(g')}_{(I)}$$
$$\underbrace{-\sum_{i\in\mathscr{I}}\sum_{g\in\Gamma_i}\alpha_i\mu_i(g)(1-\xi_i)\log\xi_i\sum_{n\geq 1}(n+1)\,\xi_i^n}_{(II)}$$
$$\underbrace{-\sum_{i\in\mathscr{I}}\alpha_i\sum_{g\in\Gamma_i}\mu_i(g)\log\mu_i(g)}_{(III)}\underbrace{-\sum_{i\in\mathscr{I}}\alpha_i\log(1-\xi_i)}_{<\infty}.$$

Convergence of (I) follows from (5.3), (II) converges since $\xi_i < 1$ and (III) is convergent by assumption $h_i < \infty$. This finishes the proof of Claim B, and thus the proof of the theorem. □

Erschler and Kaimanovich [9] asked if drift and entropy of random walks on groups depend continuously on the probability measure, which governs the random walk. Ledrappier [19] proves in his recent, simultaneous paper that drift and entropy of finite-range random walks on free groups vary analytically with the probability measure of constant support. By Theorem 5.1, we are even able to show continuity for free products of finitely generated groups, but restricted to nearest neighbour random walks with fixed set of generators.

Corollary 5.2. *Let Γ_i be generated as a semigroup by S_i. Denote by $\mathscr{P}_i$ the set of probability measures μ_i on S_i with $\mu_i(x_i) > 0$ for all $x_i \in S_i$. Furthermore, we write $\mathscr{A} := \{(\alpha_1,\dots,\alpha_r) \mid \alpha_i > 0, \sum_{i\in\mathscr{I}}\alpha_i = 1\}$. Then the entropy function*

$$h : \mathscr{A}\times\mathscr{P}_1\times\cdots\times\mathscr{P}_r \to \mathbb{R} : (\alpha_1,\dots,\alpha_r,\mu_1,\dots,\mu_r) \mapsto h(\alpha_1,\dots,\alpha_r,\mu_1,\dots,\mu_r)$$

is real-analytic.

Proof. The claim follows directly with the formula given in Theorem 5.1: the involved generating functions $F_i(o_i, g|z)$ and $G_i(o_i, o_i|z)$ are analytic when varying the probability measure of constant support, and the values ξ_i can also be rewritten as

$$\xi_i = \sum_{k_1,\dots,k_r,l_{1,1},\dots,l_{r,|S_r|}\geq 1} x(k_1,\dots,k_r,l_{1,1},\dots,l_{r,|S_r|})\prod_{i\in\mathscr{I}}\alpha_i^{k_i}\prod_{j=1}^{|S_i|}\mu_i(x_{i,j})^{l_{i,j}},$$

where $S_i = \{x_{i,1},\dots,x_{i,|S_i|}\}$. This yields the claim. □

Remarks:

1. Corollary 5.2 holds also for the case of free products of *finite* graphs, if one varies the transition probabilities continously under the assumption that the sets $\{(x_i, y_i) \in V_i \times V_i \mid p_i(x_i, y_i) > 0\}$ remain constant: one has to rewrite ξ_i as power series in terms of (finitely many) $p_i(x_i, y_i)$ and gets analyticity with the formula given in Theorem 3.7.

2. Analyticity holds also for the drift (w.r.t. the block length and w.r.t. the natural graph metric) of nearest neighbour random walks due to the formulas given in [11, Section 5 and 7].

3. The formula for entropy and drift given in Mairesse and Mathéus [21] for random walks on free products of *finite* groups depends also analytically on the transition probabilities.

6 Entropy Inequalities

In this section we consider the case of free products of *finite* sets $V_1, \dots, V_r$, where V_i has $|V_i|$ vertices. We want to establish a connection between asymptotic entropy, rate of escape and the volume growth rate of the free product V. For $n \in \mathbb{N}_0$, let $S_0(n)$ be the set of all words of V of block length n. The following lemmas give an answer how fast the free product grows.

Lemma 6.1. *The sphere growth rate w.r.t. the block length is given by*

$$s_0 := \lim_{n\to\infty} \frac{\log |S_0(n)|}{n} = \log \lambda_0,$$

where λ_0 is the Perron-Frobenius eigenvalue of the $r \times r$-matrix $D = (d_{i,j})_{1\leq,i,j\leq r}$ with $d_{i,j} = 0$ for $i = j$ and $d_{i,j} = |V_j| - 1$ otherwise.

Proof. Denote by $\widehat{D}$ the $r \times r$-diagonal matrix, which has entries $|V_1| - 1, |V_2| - 1, \dots, |V_r| - 1$ on its diagonal. Let $\mathbb{1}$ be the $(r \times 1)$-vector with all entries equal to 1. Thus, we can write

$$|S_0(n)| = \mathbb{1}^T \widehat{D} D^{n-1} \mathbb{1}.$$

Let $0 < v_1 \leq \mathbb{1}$ and $v_2 \geq \mathbb{1}$ be eigenvectors of D w.r.t. the Perron-Frobenius eigenvalue λ_0. Then

$$\begin{aligned} |S_0(n)| &\geq \mathbb{1}^T \widehat{D} D^{n-1} v_1 = C_1 \cdot \lambda_0^{n-1}, \\ |S_0(n)| &\leq \mathbb{1}^T \widehat{D} D^{n-1} v_2 = C_2 \cdot \lambda_0^{n-1}, \end{aligned}$$

where C_1, C_2 are some constants independent from n. Thus,

$$\frac{\log |S_0(n)|}{n} = \log |S_0(n)|^{1/n} \xrightarrow{n\to\infty} \log \lambda_0.$$

□

Recall from the Perron-Frobenius theorem that $\lambda_0 \geq \sum_{i=1,i\neq j}^r (|\Gamma_i| - 1)$ for each $j \in \mathscr{I}$; in particular, $\lambda_0 \geq 1$. We also take a look on the natural graph metric and its growth rate. For this purpose, we define

$$S_1(n) := \left\{x \in V \,\middle|\, p^{(n)}(o,x) > 0, \forall m < n : p^{(m)}(o,x) = 0\right\},$$

that is, the set of all vertices in V which are at distance n to the root o w.r.t. the natural graph metric.

We now construct a new graph, whose adjacency matrix allows us to describe the exponential growth of $S_1(n)$ as $n \to \infty$. For this purpose, we visualize the sets $V_1, \dots, V_r$ as graphs $\mathscr{X}_1, \dots, \mathscr{X}_r$ with vertex sets $V_1, \dots, V_r$ equipped with the following edges: for $x, y \in V_i$, there is a directed edge from x to y if and only if $p_i(x,y) > 0$. Consider now directed spanning trees $\mathscr{T}_1, \dots, \mathscr{T}_r$ of the graphs $\mathscr{X}_1, \dots, \mathscr{X}_r$

such that the graph distances of vertices in $\mathcal{T}_i$ to the root o_i remain the same as in $\mathcal{X}_i$. We now investigate the free product $\mathcal{T} = \mathcal{T}_1 * \cdots * \mathcal{T}_r$, which is again a tree. We make the crucial observation that $\mathcal{T}$ can be seen as the directed cover of a finite directed graph F, where F is defined in the following way:

1. The vertex set of F is given by $\{o\} \cup \bigcup_{i\in\mathcal{I}} V_i^\times$ with root o.
2. The edges of F are given as follows: first, we add all edges inherited from one of the trees $\mathcal{T}_1, \dots, \mathcal{T}_r$, where o plays the role of o_i for each $i \in \mathcal{I}$. Secondly, we add for all $i \in \mathcal{I}$ and every $x \in V_i^\times$ an edge from x to each $y \in V_j^\times$, $j \neq i$, whenever there is an edge from o_i to y in $\mathcal{T}_j$.

The tree $\mathcal{T}$ can be seen as a *periodic tree*, which is also called a *tree with finitely many cone types*; for more details we refer to Lyons [20] and Nagnibeda and Woess [23]. Now we are able to state the following lemma:

Lemma 6.2. *The sphere growth rate w.r.t. the natural graph metric defined by*

$$s_1 := \lim_{n\to\infty} \frac{\log |S_1(n)|}{n}$$

exists. Moreover, we have the equation $s_1 = \log \lambda_1$, where λ_1 is the Perron-Frobenius eigenvalue of the adjacency matrix of the graph F.

Proof. Since the graph metric remains invariant under the restriction of V to $\mathcal{T}$ and since it is well-known that the growth rate exists for periodic trees (see Lyons [20, Chapter 3.3]), we have existence of the limit s_1. More precisely, $|S_1(n)|^{1/n}$ tends to the Perron-Frobenius eigenvalue of the adjacency matrix of F as $n \to \infty$. For sake of completeness, we remark that the root of $\mathcal{T}$ plays a special role (as a cone type) but this does not affect the application of the results about directed covers to our case. □

For $i \in \{0,1\}$, we write $B_i(n) = \bigcup_{k=0}^n S_i(k)$. Now we can prove:

Lemma 6.3. *The volume growth w.r.t. the block length, w.r.t. the natural graph metric respectively, is given by*

$$g_0 := \lim_{n\to\infty} \frac{\log |B_0(n)|}{n} = \log \lambda_0, \quad g_1 := \lim_{n\to\infty} \frac{\log |B_1(n)|}{n} = \log \lambda_1 \quad \textit{respectively}.$$

Proof. For ease of better readability, we omit the subindex $i \in \{0,1\}$ in the following, since the proofs for g_0 and g_1 are completely analogous. Choose any small $\varepsilon > 0$. Then there is some K_ε such that for all $k \geq K_\varepsilon$

$$\lambda^k e^{-k\varepsilon} \leq |S(k)| \leq \lambda^k e^{k\varepsilon}.$$

Write $C_\varepsilon = \sum_{i=0}^{K_\varepsilon - 1} |S(i)|$. Then for $n \geq K_\varepsilon$:

$$\begin{aligned} |B(n)|^{1/n} &= \sqrt[n]{\sum_{k=0}^n |S(k)|} \leq \sqrt[n]{C_\varepsilon + \sum_{k=K_\varepsilon}^n \lambda^k e^{k\varepsilon}} = \lambda e^\varepsilon \sqrt[n]{\frac{C_\varepsilon}{\lambda^n e^{n\varepsilon}} + \sum_{k=K_\varepsilon}^n \frac{1}{\lambda^{n-k} e^{(n-k)\varepsilon}}} \\ &\leq \lambda e^\varepsilon \sqrt[n]{\frac{C_\varepsilon}{\lambda^n e^{n\varepsilon}} + (n - K_\varepsilon + 1)} \xrightarrow{n\to\infty} \lambda e^\varepsilon. \end{aligned}$$

In the last inequality we used the fact $\lambda \geq 1$. Since we can choose $\varepsilon > 0$ arbitrarily small, we get $\limsup_{n\to\infty} |B(n)|^{1/n} \leq \lambda$. Analogously:

$$|B(n)|^{1/n} \geq \sqrt[n]{C_\varepsilon + \sum_{k=K_\varepsilon}^{n} \lambda^k e^{-k\varepsilon}} = \lambda \sqrt[n]{\frac{C_\varepsilon}{\lambda^n} + \sum_{k=K_\varepsilon}^{n} \frac{e^{-k\varepsilon}}{\lambda^{n-k}}} \xrightarrow{n\to\infty} \lambda e^{-\varepsilon}.$$

That is, $\lim_{n\to\infty} \frac{1}{n} \log |B(n)| = \log \lambda$. □

For $i \in \{0,1\}$, define $l_i : V \to \mathbb{N}_0$ by $l_0(x) = \|x\|$ and $l_1(x) = \inf\{m \in \mathbb{N}_0 \mid p^{(m)}(o,x) > 0\}$. Then the limits $\ell_i = \lim_{n\to\infty} l_i(X_n)/n$ exist; see [11, Theorem 3.3, Section 7.II]. Now we can establish a connection between entropy, rate of escape and volume growth:

Corollary 6.4. *$h \leq g_0 \cdot \ell_0$ and $h \leq g_1 \cdot \ell_1$.*

Proof. Let be $i \in \{0,1\}$ and $\varepsilon > 0$. Then there is some $N_\varepsilon \in \mathbb{N}$ such that for all $n \geq N_\varepsilon$

$$1-\varepsilon \leq \mathbb{P}\Big(\big\{x \in V \mid -\log \pi_n(x) \geq (h-\varepsilon)n, l_i(x) \leq (\ell_i+\varepsilon)n\big\}\Big) \leq e^{-(h-\varepsilon)n} \cdot \Big|B_i\big((\ell_i+\varepsilon)n\big)\Big|.$$

That is,

$$(h-\varepsilon) + \frac{\log(1-\varepsilon)}{n} \leq (\ell_i+\varepsilon) \cdot \frac{\log\Big|B_i\big((\ell_i+\varepsilon)n\big)\Big|}{(\ell_i+\varepsilon)n}.$$

If we let n tend to infinity and make ε arbitrarily small, we get the claim. □

Finally, we remark that an analogous inequality for random walks on groups was given by Guivarc'h [13], and more generally for space- and time-homogeneous Markov chains by Kaimanovich and Woess [15, Theorem 5.3].

7 Examples

7.1 Free Product of Finite Graphs

Consider the graphs $\mathscr{X}_1$ and $\mathscr{X}_2$ with the transition probabilities sketched in Figure 7.1. We set $\alpha_1 = \alpha_2 = 1/2$. For the computation of ℓ_0 we need the following functions:

$$F_1(g_1,o_1|z) = \frac{z^2}{2}\frac{1}{1-z^2/2}, \qquad F_2(h_1,o_2|z) = \frac{z^2}{2}\frac{1}{1-z^3/2},$$

$$\xi_1(z) = \frac{z/2}{1-\frac{z}{2}\frac{\xi_2(z)^2}{2}\frac{1}{1-\xi_2(z)^3/2}}, \qquad \xi_2(z) = \frac{z/2}{1-\frac{z}{2}\frac{\xi_1(z)^2}{2}\frac{1}{1-\xi_1(z)^2/2}}.$$

Simple computations with the help of [11, Section 3] and Mathematica allow us to determine the rate of escape of the random walk on $\mathscr{X}_1 * \mathscr{X}_2$ as $\ell_0 = 0.41563$. For the computation of the entropy, we need also the following generating functions:

$$L_1(o_1,g_1|z) = \frac{z}{1-z^2/2}, \quad L_1(o_1,g_2|z) = \frac{z^2}{1-z^2/2}, \quad L_2(o_2,h_1|z) = \frac{z}{1-z^3/2},$$

$$L_2(o_2,h_2|z) = \frac{z^2}{1-z^3/2}, \quad L_2(o_2,h_3|z) = \frac{z^3/2}{1-z^3/2}.$$

Thus, we get the asymptotic entropy as $h = 0.32005$.

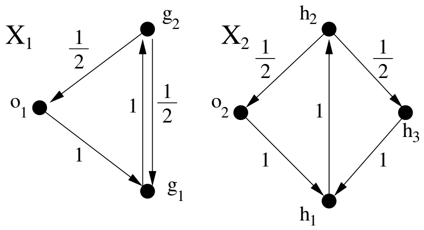


Figure 1: Finite graphs $\mathscr{X}_1$ and $\mathscr{X}_2$

7.2 $(\mathbb{Z}\times\mathbb{Z}/2)*(\mathbb{Z}\times\mathbb{Z}/2)$

Consider the free product $\Gamma=\Gamma_1*\Gamma_2$ of the *infinite* groups $\Gamma_i=\mathbb{Z}\times(\mathbb{Z}/2\mathbb{Z})$ with $\alpha_i=1/2$ and $\mu_i\big((\pm1,0)\big)=\mu_i\big((0,1)\big)=1/3$ for each $i\in\{1,2\}$. We set $a:=(1,0)$, $b:=(1,1)$, $c:=(0,1)$ and $\lambda(x,y):=x$ for $(x,y)\in\Gamma_i$. Define

$$\begin{aligned}\hat{F}(a|z) &:= \sum_{n\geq1}\mathbb{P}\Big[\,Y_n=a,\forall m<n:\lambda(Y_m)<1\,\Big|\,Y_0=(0,0)\Big]\,z^n,\\ \hat{F}(b|z) &:= \sum_{n\geq1}\mathbb{P}\Big[\,Y_n=b,\forall m<n:\lambda(Y_m)<1\,\Big|\,Y_0=(0,0)\Big]\,z^n,\end{aligned}$$

where $(Y_n)_{n\in\mathbb{N}_0}$ is a random walk on $\mathbb{Z}\times\mathbb{Z}/2$ governed by μ_1. The above functions satisfy the following system of equations:

$$\begin{aligned}\hat{F}(a|z) &= \frac{z}{3}\Big(1+\hat{F}(b|z)+\hat{F}(a|z)^2+\hat{F}(b|z)^2\Big),\\ \hat{F}(b|z) &= \frac{z}{3}\Big(\hat{F}(a|z)+\hat{F}(a|z)\hat{F}(b|z)+\hat{F}(b|z)\hat{F}(a|z)\Big).\end{aligned}$$

From this system we obtain explicit formulas for $\hat{F}(a|z)$ and $\hat{F}(b|z)$. We write $F\big(n,j|z\big):=F_1\big((0,0),(n,j)|z\big)$ for $(n,j)\in\mathbb{Z}\times\mathbb{Z}/2$. To compute the entropy rate we have to solve the following system of equations:

$$\begin{aligned}F(a|z) &= \frac{z}{3}\Big(1+F(b|z)+\hat{F}(a|z)F(a|z)+\hat{F}(b|z)F(b|z)\Big),\\ F(b|z) &= \frac{z}{3}\Big(F(c|z)+F(a|z)+\hat{F}(a|z)F(b|z)+\hat{F}(b|z)F(a|z)\Big),\\ F(c|z) &= \frac{z}{3}\Big(1+2\,F(b|z)\Big).\end{aligned}$$

Moreover, we need the value $\xi_1(1)=\xi_2(1)=\xi$. This value can be computed analogously to [11, Section 6.2], that is, ξ has to be computed numerically from the equation

$$\frac{\xi}{2-2\xi}=\xi\,G_1(\xi)=\frac{\xi}{1-\frac{2}{3}\xi F(a|\xi)-\frac{1}{3}\xi F(c|\xi)}.$$

Solving this equation with MATHEMATICA yields $\xi=0.55973$. To compute the entropy we have to evaluate the functions $F(g|z)$ at $z=\xi$ for each $g\in\mathbb{Z}\times\mathbb{Z}_2$. For even $n\in\mathbb{N}$, we have the following formulas:

$$
\begin{aligned}
F\big((\pm n,0)|\xi\big) &= \sum_{k=0}^{n/2} \binom{n}{2k} \hat{F}(b|\xi)^{2k}\hat{F}(a|\xi)^{n-2k} + \\
&\quad \sum_{k=0}^{n/2-1} \binom{n}{2k+1} \hat{F}(b|\xi)^{2k+1}\hat{F}(a|\xi)^{n-2k-1}F(c|\xi), \\
F\big((\pm n,1)|\xi\big) &= \sum_{k=0}^{n/2-1} \binom{n}{2k+1} \hat{F}(b|\xi)^{2k+1}\hat{F}(a|\xi)^{n-2k-1} + \\
&\quad \sum_{k=0}^{n/2} \binom{n}{2k} \hat{F}(b|\xi)^{2k}\hat{F}(a|\xi)^{n-2k}F(c|\xi).
\end{aligned}
$$

For odd $n \in \mathbb{N}$,

$$
\begin{aligned}
F\big((\pm n,0)|\xi\big) &= \sum_{k=0}^{(n-1)/2} \binom{n}{2k} \hat{F}(b|\xi)^{2k}\hat{F}(a|\xi)^{n-2k} + \\
&\quad \sum_{k=0}^{(n-1)/2} \binom{n}{2k+1} \hat{F}(b|\xi)^{2k+1}\hat{F}(a|\xi)^{n-2k-1}F(c|\xi), \\
F\big((\pm n,1)|\xi\big) &= \sum_{k=0}^{(n-1)/2} \binom{n}{2k+1} \hat{F}(b|\xi)^{2k+1}\hat{F}(a|\xi)^{n-2k-1} + \\
&\quad \sum_{k=0}^{(n-1)/2} \binom{n}{2k} \hat{F}(b|\xi)^{2k}\hat{F}(a|\xi)^{n-2k}F(c|\xi).
\end{aligned}
$$

Moreover, we define $\hat{F} := \mathbb{P}[\exists n \in \mathbb{N} : \lambda(X_n) = 1]$. This probability can be computed by conditioning on the first step and solving

$$
\hat{F} = \frac{\xi}{3}\big(1 + \hat{F} + \hat{F}^2\big),
$$

that is, $\hat{F} = 0.24291$. Observe that we get the following estimations:

$$
\begin{aligned}
F_1(o,g|\xi) &\leq \hat{F}^{|\lambda(g)|} \quad \text{for } g \in \mathbb{Z} \times \mathbb{Z}_2, \\
F_1(o,g|\xi) &\geq \hat{F}^{|\lambda(g)|-1} \cdot \min\big\{F_1(o_1,a|\xi), F_1(o_1,b|\xi)\big\} \quad \text{for } g \in \big(\mathbb{Z} \times \mathbb{Z}_2\big) \setminus \{(0,0),c\}.
\end{aligned}
$$

These bounds allow us to cap the sum over all $g' \in \Gamma_i^{\times}$ in (5.2) and to estimate the tails of these sums. Thus, we can compute the entropy rate numerically as $h = 1.14985$.

Acknowledgements

The author is grateful to Frédéric Mathéus for discussion on the problems and several hints regarding content and exposition.

References

[1] A. Avez. Entropie des groupes de type fini. *C.R.Acad.Sci.Paris Sér. A-B*, 275:A1363–A1366, 1972. MR0324741

[2] I. Benjamini and Y. Peres. Tree-indexed random walks on groups and first passage percolation. *Probab. Theory Related Fields*, 98(1): 91–112, 1994. MR1254826

[3] S. Blachère and P. Haïssinsky and P. Mathieu. Asymptotic entropy and Green speed for random walks on countable groups. *Ann. of Probab.*, 36(3):1134–1152, 2008. MR2408585

[4] E. Candellero and L. Gilch. Phase Transitions for Random Walk Asymptotics on Free Products of Groups. Submitted, current version available at arXiv:0909.1893v3, 2009.

[5] D.I. Cartwright and P.M. Soardi. Random walks on free products, quotients and amalgams. *Nagoya Math. J.*, 102:163–180, 1986. MR0846137

[6] T. Cover and J. Thomas. *Elements of Information Theory*. Wiley & Sons, 2nd edition, 2006. MR2239987

[7] Y. Derriennic. Quelques applications du théorème ergodique sous-additif. *Astérisque*, 74:183–201, 1980. MR0588163

[8] A. Erschler. On the asymptotics of drift. *J. of Math. Sciences*, 121(3):2437–2440, 2004. MR1879073

[9] A. Erschler and V. Kaimanovich. Continuity of entropy for random walks on hyperbolic groups. *In preparation*, 2010.

[10] P. Gerl and W. Woess. Local Limits and Harmonic Functions for Nonisotropic Random Walks on Free Groups. *Probab. Theory Rel. Fields*, 71:341–355, 1986. MR0824708

[11] L. A. Gilch. Rate Of Escape of Random Walks on Free Products. *J. Aust. Math. Soc.*, 83(I):31–54, 2007. MR2378433

[12] L. A. Gilch and S. Müller. Random Walks on Directed Covers of Graphs. *J. of Theoret. Probab.*, DOI 10.1007/s10959-009-0256-0, 2009.

[13] Y. Guivarc'h. Sur la loi des grands nombres et le rayon spectral d'une marche aléatoire. *Astérisque*, 74:47–98, 1980. MR0588157

[14] V.A. Kaimanovich and A.M. Vershik. Random walks on discrete groups: boundary and entropy. *Ann. of Prob.*, 11:457–490, 1983. MR0704539

[15] V. Kaimanovich and W. Woess. Boundary and entropy of space homogeneous Markov chains. *Ann. of Prob.*, 30:323–363, 2002. MR1894110

[16] A. Karlsson and F. Ledrappier. Linear drift and Poisson boundary for random walks. *Pure Appl. Math. Q.*, 3:1027–1036, 2007. MR2402595

[17] J.F.C. Kingman. The ergodic theory of subadditive processes. *J. Royal Stat. Soc., Ser. B*, 30:499–510, 1968. MR0254907

[18] S. Lalley. Finite range random walk on free groups and homogeneous trees. *Ann. of Prob.*, 21(4):2087–2130, 1993. MR1245302

[19] F. Ledrappier. Analyticity of the entropy for some random walks. *Preprint, available at arXiv:1009.5354v1*, 2010.

[20] R. Lyons, with Y. Peres. *Probability on Trees and Networks*. In preparation. Current version available at `http://mypage.iu.edu/~rdlyons/`, 2010.

[21] J. Mairesse and F. Mathéus. Random walks on free products of cyclic groups. *J. London Math. Soc.*, 75(1):47–66, 2007. MR2302729

[22] J.C. McLaughlin. *Random Walks and Convolution Operators on Free Products.* PhD thesis, New York Univ., 1986. MR2635581

[23] T. Nagnibeda and W. Woess. Random walks on trees with finitely many cone types. *J. Theoret. Probab.*, 15:399–438, 2002. MR1898814

[24] S. Sawyer. Isotropic random walks in a tree. *Zeitschrift f. Wahrscheinlichkeitstheorie*, 42:279–292, 1978. MR0491493

[25] S. Sawyer and T. Steger. The rate of escape for anisotropic random walks in a tree. *Probab. Theory Rel. Fields*, 76:207–230, 1987. MR0906775

[26] N.T. Varopoulos. Long range estimates for Markov chains. *Bull. Sc. math.*, 109:225–252, 1985. MR0822826

[27] D. Voiculescu. Addition of certain non-commuting random variables. *J. Funct. Anal.*, 66:323–346, 1986. MR0839105

[28] W. Woess. Nearest neighbour random walks on free products of discrete groups. *Boll. Un. Mat. Ital.*, 5-B:961–982, 1986. MR0871708

[29] W. Woess. *Random Walks on Infinite Graphs and Groups*. Cambridge University Press, 2000. MR1743100

Publication B

Phase Transitions for Random Walk Asymptotics on Free Products of Groups

E. Candellero and L. Gilch

Random Structures and Algorithms,
Vol. 40, Issue 2, 150–181, 2012.

Phase Transitions for Random Walk Asymptotics on Free Products of Groups*

Elisabetta Candellero, Lorenz A. Gilch
Institut für Mathematische Strukturtheorie (Math. C), University of Technology Graz, Steyrergasse 30, A-8010 Graz, Austria; e-mail: candellero@tugraz.at; gilch@tugraz.at

Received 10 September 2009; accepted 11 February 2011; received in final form 18 March 2011
Published online 31 May 2011 in Wiley Online Library (wileyonlinelibrary.com).
DOI 10.1002/rsa.20370

ABSTRACT: Suppose we are given finitely generated groups $\Gamma_1, \ldots, \Gamma_m$ equipped with irreducible random walks. Thereby we assume that the expansions of the corresponding Green functions at their radii of convergence contain only logarithmic or algebraic terms as singular terms up to sufficiently large order (except for some degenerate cases). We consider transient random walks on the free product $\Gamma_1 * \ldots * \Gamma_m$ and give a complete classification of the possible asymptotic behaviour of the corresponding n-step return probabilities. They either inherit a law of the form $\varrho^{n\delta} n^{-\lambda_i} \log^{\kappa_i} n$ from one of the free factors Γ_i or obey a $\varrho^{n\delta} n^{-3/2}$-law, where $\varrho < 1$ is the corresponding spectral radius and δ is the period of the random walk. In addition, we determine the full range of the asymptotic behaviour in the case of nearest neighbour random walks on free products of the form $\mathbb{Z}^{d_1} * \ldots * \mathbb{Z}^{d_m}$. Moreover, we characterize the possible phase transitions of the non-exponential types $n^{-\lambda_i} \log^{\kappa_i} n$ in the case $\Gamma_1 * \Gamma_2$. © 2011 Wiley Periodicals, Inc. Random Struct. Alg., 40, 150–181, 2012

Keywords: random walks; free products; return probabilities; asymptotic behaviour; lattices

1. INTRODUCTION

In this article we investigate transient random walks on free products $\Gamma_1 * \ldots * \Gamma_m$, where $m \geq 2$ and $\Gamma_1, \ldots, \Gamma_m$ are finitely generated groups. These random walks arise from convex combinations of probability measures on the factors $\Gamma_1, \ldots, \Gamma_m$. Our aim is to compute the asymptotic behaviour of the n-step return probabilities on the free product. In a general setting, one has a typical asymptotic behaviour of the form $\mu^{(n)}(x) \sim C_x \varrho^{n\delta} n^{-\lambda}$, where $\mu^{(n)}(x)$ is the probability of being at x at time n, ϱ is the spectral radius, δ the period of the random walk, and C_x some constant depending on x. If e is the group identity and

Correspondence to: L. A. Gilch
*Supported by Research and Technology Office at TU Graz, NAWI-Graz; German Research Foundation (DFG) (GI 746/1-1); Austrian Academy of Science (ÖAW); European Science Foundation (ESF)

starting point of the random walk, then $\mu^{(n)}(e)$ is called the n-step return probability. Gerl [8] conjectured that the n-step return probabilities of two symmetric measures on a group satisfying such a limit law have the same $n^{-\lambda}$, that is, λ is a group invariant. Cartwright [2] came to the astonishing result that for random walks on $\mathbb{Z}^d * \mathbb{Z}^d$ with $d \geq 5$ there are at least two possible types of asymptotic behaviour, namely $n^{-3/2}$ and $n^{-d/2}$. In relation with his joint work with Chatterji et al. [5] L. Saloff-Coste asked whether the range of different asymptotic behaviours can still be wider than in the case considered by Cartwright. In this article we will pick up this question by investigating more general laws of the form $C\,\varrho^{n\delta}\,n^{-\lambda}\,\log^{\kappa} n$. In this case, we speak of the factor $n^{-\lambda}\,\log^{\kappa} n$ as the non-exponential type of the return probabilities.

The starting point for the present investigation was Woess [22, Section 17.B], where the result of Cartwright [2] is explained that simple random walk on $\mathbb{Z}^d * \mathbb{Z}^d$ for $d \geq 5$ satisfies a $n^{-d/2}$-law. In this article we will prove the following more general theorem:

Theorem 1.1. *Let $m \in \mathbb{N}$ with $m \geq 2$ and $d_1, \dots, d_m \in \mathbb{N}$. For each $i \in \{1, \dots, m\}$, consider on the lattice $\mathbb{Z}^{d_i}$ a probability measure μ_i with $\operatorname{supp}(\mu_i) = \{\pm e_j^{(i)} \mid 1 \leq j \leq d_i\}$, where $e_j^{(i)}$ is the j-th unit vector in $\mathbb{Z}^{d_i}$. For any $\alpha_1, \dots, \alpha_m > 0$ with $\sum_{i=1}^m \alpha_i = 1$, let $\mu := \sum_{i=1}^m \alpha_i \mu_i$ govern a (irreducible) random walk on the free product $\mathbb{Z}^{d_1} * \dots * \mathbb{Z}^{d_m}$ starting at e, where e denotes the identity of the free product.*

Then the return probabilities $\mu^{(2n)}(e)$ behave asymptotically either like $C \cdot \varrho^{2n} \cdot n^{-d_i/2}$ for $i \in \{1, \dots, m\}$ or like $C \cdot \varrho^{2n} \cdot n^{-3/2}$ for some constant $C = C_\mu$ depending on μ. Moreover, if all exponents d_i are different and $\min\{d_1, \dots, d_m\} \geq 5$ then exactly $m+1$ different asymptotic behaviours may occur by choosing the random walk adequately.

We will consider more general free products which go beyond free products of lattices. For this purpose, we will present a new approach in order to be able to deal with irreducible random walks on any free product of the form $\Gamma_1 * \dots * \Gamma_m$. At this point we have to make the following assumption: if the Green function of the random walk on the free factor Γ_i is differentiable at its radius of convergence $\mathbf{r}_i$, then the Green function is assumed to have a singular expansion (i.e. in a neighbourhood of $\mathbf{r}_i$) containing only singular terms of the form $(\mathbf{r}_i - z)^q \log^k(\mathbf{r}_i - z)$ with $q \in (1, \infty)$ and $k \in \mathbb{N}_0$ up to sufficiently large order. The latter property is satisfied for several well-known groups as e.g. $\mathbb{Z}^d$ or $\mathbb{Z}^d \times (\mathbb{Z}/n\mathbb{Z})$ with $d \geq 5$ and $n \geq 2$. If, however, the Green function of the random walk on the free factor Γ_i is *not* differentiable at $\mathbf{r}_i$, we do *not* need any assumption on the expansions.

If the asymptotic n-step return probabilities of the random walk on Γ_i satisfy a law of the form $\mathbf{r}_i^{-n\delta} n^{-\lambda_i} \log^{\kappa_i} n$ then we will show that only one of the following non-exponential types may occur for the random walk on the free product: either $n^{-\lambda_i} \log^{\kappa_i} n$ for some $i \in \{1, \dots, m\}$, or $n^{-3/2}$. That is, we may have up to $m+1$ different types of asymptotic behaviour for (symmetric or non-symmetric) random walks, and Theorem 1.1 shows that one can have indeed exactly $m+1$ different behaviours. Moreover, for the case $\Gamma_1 * \Gamma_2$ equipped with the probability measure $\mu = \alpha_1\mu_1 + (1-\alpha_1)\mu_2$, where μ_1 and μ_2 are probability measures on Γ_1 and Γ_2 and $\alpha_1 \in (0,1)$, we characterize the phase transitions of the non-exponential types in terms of α_1. We split the $(0,1)$-interval, i.e. the interval of possible values for α_1, in up to three distinct subintervals such that, in each of them, we have exactly one of the non-exponential types $n^{-\lambda_1} \log^{\kappa_1}(n)$, $n^{-\lambda_2} \log^{\kappa_2} n$ or $n^{-3/2}$.

Let us briefly recall some results regarding the asymptotic behaviour of return probabilities. Work in this direction has been done since the 1970's by Gerl, Sawyer, Woess, Cartwright, Soardi and Lalley, see e.g. [3,9,12,18,21]. Sawyer [18] applies Fourier analysis to isotropic random walks on trees (free groups), which uses in a crucial way methods from complex analysis. For finite range random walks on free groups, it is known from [9] and [12] that the n-step return probabilities behave asymptotically like $C\varrho^n n^{-3/2}$, where $\varrho < 1$. In [8,20,21] free products of finite groups are considered, which have a very tree-like structure and where random walks obey a $n^{-3/2}$-law. In the more general case of free products of arbitrary groups the interior structure of each free factor is more complicated. Woess [21], Cartwright and Soardi [3], Voiculescu [19] and McLaughlin [16] found independently a method to determine the Green function of a free product in terms of functional equations of the Green functions defined on the free factors. We will study these equations carefully, in order to obtain – with the help of the well-known *method of Darboux* – the asymptotic behaviour of the power series' coefficients, which are the sought return probabilities. We refer also to the survey of Woess [23], which outlines the use of generating functions. More recently, random walks on free products have also been studied by Mairesse and Mathéus [15] and Gilch [10, 11], regarding boundary theory, entropy and rate of escape. For more details and references we refer to Woess [22], which serves also as reference text for our work.

The structure of this paper is as follows: in Section 2 we introduce some basic facts and notations. In Section 3 we prove our main result for the case $\Gamma_1 * \Gamma_2$, while in Sections 4 and 5 we are completing the list of degenerate cases, which, in particular, may occur if the Green functions of the random walks on the single factors are non-differentiable at their radii of convergence. In Section 5.3 we are proving inductively the proposed asymptotic behaviour for multi-factor free products of the form $\Gamma_1 * \ldots * \Gamma_m$ with $m \geq 3$. Section 6 discusses some examples. This includes the case of free products of the form $\mathbb{Z}^{d_1} * \ldots * \mathbb{Z}^{d_m}$, where we give a full classification of the asymptotic behaviour of the return probabilities, which proves Theorem 1.1. For $\Gamma_1 * \Gamma_2$, we give in Section 7 a full characterization of the possible phase transition behaviour of the non-exponential types of the return probabilities in terms of the weight α_1 of the probability measure given on Γ_1. Finally, Section 8 gives some concluding remarks about higher asymptotic order terms.

2. RANDOM WALKS ON FREE PRODUCTS

Let $m \in \mathbb{N}$ with $m \geq 2$. Suppose we are given finitely generated groups $\Gamma_1, \ldots, \Gamma_m$, and we denote by e_i the identity of Γ_i. We consider the *free product* $\Gamma := \Gamma_1 * \ldots * \Gamma_m$, which consists of all finite words of the form

$$x_1 x_2 \ldots x_n, \tag{2.1}$$

where $x_1, \ldots, x_n \in \bigcup_{i=1}^m \Gamma_i \setminus \{e_i\}$ and two consecutive letters are not from the same free factor Γ_i. In the case $\Gamma_i = \Gamma_j$ we may think that the elements of Γ_i and Γ_j have different colours to distinguish their origin. Observe that each factor Γ_i can be naturally embedded into Γ, and therefore $e_i \in \Gamma_i$ can be identified with the empty word $e \in \Gamma$. The free product is a group with e as identity: the product of two elements is given by concatenation followed by iterated contractions and cancellations of redundant terms in the middle, in order to obtain the requested form (2.1). For example, if $a, b \in \Gamma_1 \setminus \{e_1\}$ and $c \in \Gamma_2 \setminus \{e_2\}$, such that $c^2 \neq e$, then $(aca^{-1})(aca) = ac^2a$. For further details about free products see e.g. Lyndon and Schupp [14].

We recall and introduce some notation: for any function $f : D \subseteq \mathbb{C} \to \mathbb{C}$ with $f(z_0) = 0$ for $z_0 \in D$, $0 < q \in \mathbb{R}$ and $k \in \mathbb{N}_0$, we use the notation $f(z) = \mathbf{o}((z_0 - z)^q \log^k(z_0 - z))$,

$f(z) = \mathcal{O}_c((z_0 - z)^q \log^k(z_0 - z))$ or $f(z) = \mathcal{O}((z_0 - z)^q \log^k(z_0 - z))$, if for $z \to z_0$ the function $f(z)$ divided by $(z_0 - z)^q \log^k(z_0 - z)$ tends to zero, has a non-zero finite limit or is bounded nearby z_0 (that is, the quotient has a finite limes superior) respectively. Furthermore, we write $(z_0 - z)^{q_1} \log^{k_1}(z_0 - z) \preceq (z_0 - z)^{q_2} \log^{k_2}(z_0 - z)$ if and only if $(z_0 - z)^{q_2} \log^{k_2}(z_0 - z) = \mathcal{O}((z_0 - z)^{q_1} \log^{k_1}(z_0 - z))$. The value z_0 will be obvious from the context.

Suppose we are given probability measures μ_i on Γ_i with $\langle \mathrm{supp}(\mu_i)\rangle = \Gamma_i$ for each $i \in \{1, \dots, m\}$. These measures μ_i govern random walks on Γ_i, that is, the single step transition probabilities are given by $p_i(x_i, y_i) = \mu_i(x_i^{-1}y_i)$ for all $x_i, y_i \in \Gamma_i$. (Let us remark that, in the case of $\Gamma_i = \mathbb{Z}^{d_i}$ with $d_i \in \mathbb{N}$, we speak of a nearest neighbour random walk if $\mathrm{supp}(\mu_i) = \{\pm e_j \mid 1 \leq j \leq d_i\}$, where e_j is the j-th unit vector in $\mathbb{Z}^{d_i}$.) We lift now μ_i to a probability measure $\bar{\mu}_i$ on Γ by defining $\bar{\mu}_i(x) := \mu_i(x)$ if $x \in \Gamma_i$; otherwise we set $\bar{\mu}_i(x) := 0$. Let $\alpha_1, \dots, \alpha_m > 0$ such that $\sum_{i=1}^m \alpha_i = 1$. Consider now the probability measure $\mu := \sum_{i=1}^m \alpha_i \bar{\mu}_i$ on the free product Γ, which arises as a convex combination of the $\bar{\mu}_i$'s. Then the single step transition probabilities on Γ given by $p(x, y) := \mu(x^{-1}y)$ for $x, y \in \Gamma$ define a random walk on Γ, which is an irreducible Markov chain. We denote by $\mu_1^{(n)}, \dots, \mu_m^{(n)}$ and $\mu^{(n)}$ the n-fold convolution power of $\mu_1, \dots, \mu_m$ and μ, that is, the distribution after n steps with start at the identity. For $z \in \mathbb{C}$, the associated *Green functions* of the random walks on Γ_i and Γ are given by

$$G_i(z) := \sum_{n=0}^{\infty} \mu_i^{(n)}(e_i) z^n \quad \text{and} \quad G(z) := \sum_{n=0}^{\infty} \mu^{(n)}(e) z^n.$$

The corresponding radii of convergence are denoted by $\mathbf{r}_i$ and $\mathbf{r}$ respectively, which are singularities according to Pringsheim's Theorem. Note that $\mathbf{r} > 1$, since Γ is non-amenable unless $\Gamma = (\mathbb{Z}/2\mathbb{Z}) * (\mathbb{Z}/2\mathbb{Z})$ (see e.g. [22, Theorem 10.10, Corollary 12.5]; in the latter case the random walk on Γ is recurrent). In the following we assume that $G_i(z)$ is exactly d_i-times differentiable at $z = \mathbf{r}_i$, where $d_i \in \mathbb{N}_0$. At this point we make the *basic assumption* that – whenever $G_i'(\mathbf{r}_i) < \infty$ – the expansions of the Green functions $G_i(z)$ in a neighbourhood of $z = \mathbf{r}_i$ have the form

$$G_i(z) = \sum_{k=0}^{d_i} g_k^{(i)} (\mathbf{r}_i - z)^k + \sum_{(q,k) \in \mathcal{T}_i} g_{(q,k)}^{(i)} (\mathbf{r}_i - z)^q \log^k(\mathbf{r}_i - z) + \mathcal{O}((\mathbf{r}_i - z)^{d_i + 2}), \qquad (2.2)$$

where $\mathcal{T}_i$ is a finite subset of $\{(q, k) \in \mathbb{R} \times \mathbb{N}_0 \mid d_i < q \leq d_i + 2\}$. In other words, the expansions contain only logarithmic and algebraic terms as singular terms up to order $(\mathbf{r}_i - z)^{d_i+2}$. As we will see, higher order terms are not necessary for the computation of the non-exponential type of the n-step return probabilities of the random walk on Γ. In the following we want to motivate this assumption on $G_i(z)$. This property for the expansion is satisfied in several well-known cases: for example, the Green functions of nearest neighbour random walks on lattices $\mathbb{Z}^d$ have such an expansion; see Proposition 6.1. With some effort, such an expansion can be deduced for $\mathbb{Z}^d \times (\mathbb{Z}/n\mathbb{Z})$ via the same methods used for $\mathbb{Z}^d$. In particular, we will prove our main result by induction on the number m of free factors of Γ: we will see that the assumptions stated in (2.2) are stable under free products (except for some degenerate cases), that is, $G(z)$ has again a similar expansion if $G'(\mathbf{r}) < \infty$ holds. If the Green function $G_i(z)$ has the form (2.2) then the well-known *method of Darboux* yields that the n-step return probabilities of the random walk on Γ_i (governed by μ_i) behave asymptotically like the coefficients of the Taylor expansion of the leading singular term in

(2.2) in a neighbourhood of 0. Assume that $S_i(z) := (\mathbf{r}_i - z)^{q_i} \log^{k_i}(\mathbf{r}_i - z)$ is the *smallest* (or *leading*) singular term in (2.2) w.r.t. $\preceq$, that is, $q > q_i$ or $(q = q_i \wedge k < k_i)$ for all $(q,k) \in \mathcal{T}_i \setminus \{(q_i,k_i)\}$; then the coefficients of its expansion in a neighbourhood of 0 behave asymptotically like the n-step return probabilities on Γ_i (the proof of this fact is completely analogous to the one of Theorem 3.1). More precisely, they behave like $\hat{C}_i \mathbf{r}_i^{-n\delta_i} n^{-\lambda_i} \log^{\kappa_i}(n)$, where $\delta_i := \gcd\{n \in \mathbb{N} \mid \mu_i^{(n)}(e_i) > 0\}$ is the period of the random walk on Γ_i and

$$\lambda_i := q_i + 1 \text{ and } \kappa_i := \begin{cases} k_i, & \text{if } q_i \notin \mathbb{N}, \\ k_i - 1 & \text{if } q_i \in \mathbb{N}; \end{cases} \tag{2.3}$$

see e.g. Flajolet and Sedgewick [7, Chapter VI.2] for the asymptotic behaviour of the coefficients in the expansion of $(\mathbf{r}_i - z)^{q_i} \log^{k_i}(\mathbf{r}_i - z)$ in a neighbourhood of 0. Analogously, $\delta := \gcd\{n \in \mathbb{N} \mid \mu^{(n)}(e) > 0\} = \gcd\{\delta_1, \dots, \delta_m\}$ is the period of the random walk on Γ. Note that the method of Darboux needs some differentiability assumptions at $z = \mathbf{r}_i$; therefore, we need the expansions of $G_i(z)$ up to terms of order $(\mathbf{r}_i - z)^{d_i+2}$. For more details about Darboux's method we refer to the comments in the proof of Theorem 3.1. We remark that another – modern – tool to handle singular expansions as in (2.2) is *Singularity Analysis*, which was developed by Flajolet and Odlyzko [6]. However, in our context it turns out that the verification of the specific requirements of singularity analysis is quite cumbersome as one can also see in Lalley [13]. Let us also point out that, in the case $G_i'(\mathbf{r}_i) = \infty$, we do *not* need any assumptions on the singularity type at $z = \mathbf{r}_i$.

In the following we look at free products of the form $\Gamma_1 * \Gamma_2$ different from $(\mathbb{Z}/2\mathbb{Z}) * (\mathbb{Z}/2\mathbb{Z})$ (it is well-known that random walks – in our context – on $(\mathbb{Z}/2\mathbb{Z})*(\mathbb{Z}/2\mathbb{Z})$ obey a $n^{-1/2}$-law). Free products with more than two factors are discussed in Section 5.3. We introduce the following *first visit generating functions* for $z \in \mathbb{C}$, $i \in \{1,2\}$ and all $s_i \in \text{supp}(\mu_i)$, $s \in \text{supp}(\mu) = \text{supp}(\mu_1) \cup \text{supp}(\mu_2)$:

$$F_i(s_i|z) := \sum_{n \geq 0} \mathbb{P}\big[X_n^{(i)} = e_i, \forall m < n : X_m^{(i)} \neq e_i \,\big|\, X_0^{(i)} = s_i\big]\, z^n,$$
$$F(s|z) := \sum_{n \geq 0} \mathbb{P}[X_n = e, \forall m < n : X_m \neq e \mid X_0 = s]\, z^n,$$

where $(X_n^{(i)})_{n \in \mathbb{N}_0}$ is a random walk on Γ_i governed by μ_i. By conditioning on the number of visits of e_i the functions $F_i(s_i|z)$ are directly linked with $G_i(z)$ via

$$G_i(z) = \frac{1}{1 - \sum_{s_i \in \text{supp}(\mu_i)} \mu_i(s_i)\, z\, F_i(s_i|z)}. \tag{2.4}$$

In the following we will summarize some further important basic facts, where we will refer to Woess [22] for further details. Define

$$\zeta_1(z) := \frac{\alpha_1 z}{1 - \alpha_2 z \sum_{s_2 \in \text{supp}(\mu_2)} \mu_2(s_2) F(s_2|z)} \text{ and}$$
$$\zeta_2(z) := \frac{\alpha_2 z}{1 - \alpha_1 z \sum_{s_1 \in \text{supp}(\mu_1)} \mu_1(s_1) F(s_1|z)}. \tag{2.5}$$

Note that $\zeta_i(1)$ is the probability of starting at e and making a step from e w.r.t. μ_i after finite time. Observe that $F(s_i|z) = F_i(s_i|\zeta_i(z))$ for $s_i \in \text{supp}(\mu_i)$; see [22, Proposition 9.18c)]. By

[22, Equation (9.20)] and (2.4), the functions $F_i(s_i|\zeta_i(z))$, $G_i(z)$ and $G(z)$ can be linked as follows:

$$G(z) = \frac{\zeta_i(z)}{\alpha_i z} G_i(\zeta_i(z)) = \frac{\zeta_i(z)}{\alpha_i z \left(1 - \sum_{s_i \in \mathrm{supp}(\mu_i)} \mu_i(s_i)\, \zeta_i(z)\, F_i(s_i|\zeta_i(z))\right)}. \tag{2.6}$$

Hence, our aim will be to determine an expansion of $\zeta_i(z)$ in a neighbourhood of $z = \mathbf{r}$, in order to get a singular expansion for $G(z)$ in a neighbourhood of $z = \mathbf{r}$. By [22, Proposition 9.10], there are functions Φ_i, $i \in \{1, 2\}$, and Φ such that

$$G_i(z) = \Phi_i(zG_i(z)) \text{ and } G(z) = \Phi(zG(z)) \tag{2.7}$$

for all $z \in \mathbb{C}$ in an open neighbourhood of the intervals $[0, \mathbf{r}_i)$ and $[0, \mathbf{r})$ respectively. In particular, the functions Φ_i and Φ are analytic in an open neighbourhood of the intervals $[0, \theta_i)$ and $[0, \theta)$ respectively, where $\theta_i := \mathbf{r}_i G_i(\mathbf{r}_i)$ and $\theta := \mathbf{r}G(\mathbf{r})$. Φ_i and Φ are also strictly increasing and strictly convex in $[0, \theta_i)$ and $[0, \theta)$ respectively. Furthermore, we define

$$\Psi_i(t) := \Phi_i(t) - t\Phi_i'(t) \quad \text{and} \quad \Psi(t) := \Phi(t) - t\Phi'(t). \tag{2.8}$$

By [22, Theorem 9.19],

$$\Phi(t) = \Phi_1(\alpha_1 t) + \Phi_2(\alpha_2 t) - 1 \quad \text{and} \quad \Psi(t) = \Psi_1(\alpha_1 t) + \Psi_2(\alpha_2 t) - 1. \tag{2.9}$$

We write $\Psi_i(\theta_i) := \lim_{t \to \theta_i -} \Psi_i(t)$. Define

$$\bar{\theta} := \min\left\{\frac{\theta_1}{\alpha_1}, \frac{\theta_2}{\alpha_2}\right\}.$$

We will make a case distinction according to finiteness of $G_i(\mathbf{r}_i)$ and $G_i'(\mathbf{r}_i)$ and also to the sign of $\Psi(\bar{\theta}) := \lim_{t \to \bar{\theta}-} \Psi(t)$. If $\Psi(\bar{\theta}) < 0$ then the n-step return probabilities of the random walk on Γ behave asymptotically like

$$\mu^{(n\delta)}(e) \sim C \cdot \mathbf{r}^{-n\delta} \cdot n^{-3/2}$$

and the Green function of the random walk on Γ has the form

$$G(z) = A(z) + \sqrt{\mathbf{r} - z}B(z), \tag{2.10}$$

where $A(z), B(z)$ are analytic functions in a neighbourhood of $z = \mathbf{r}$ with $B(\mathbf{r}) \neq 0$; see [22, Theorem 17.3] or [7, Section VI.7.]. Moreover, if one fixes any finite, symmetric sets $\mathcal{S}_i$ of generators of Γ_i for $i \in \{1, 2\}$, where each $\mathcal{S}_i$ contains at least one element of order bigger than 2, then μ_1, μ_2 and α_1 can always be chosen in a suitable way in order to obtain $\Psi(\bar{\theta}) < 0$ with $\mathrm{supp}(\mu_i) = \mathcal{S}_i$; see [22, Corollary 17.10]. In particular, the same asymptotic behaviour (including an expansion of the Green function of the form (2.10)) holds if Γ_1 and Γ_2 are finite, see [21]. Therefore, we assume from now on that at least one out of Γ_1 and Γ_2 is infinite, and we may restrict our investigation to the cases $\Psi(\bar{\theta}) > 0$ and $\Psi(\bar{\theta}) = 0$.

We remark some important facts for the case $\Psi(\bar{\theta}) \geq 0$. If the latter holds, we have $\theta = \bar{\theta}$ and $G(\mathbf{r}) < \infty$, see [22, Theorem 9.22]. By [22, Lemma 17.1.a)], we have $\zeta_i(\mathbf{r}) \leq \mathbf{r}_i$ for $i \in \{1, 2\}$ with equality if and only if $\theta = \theta_i/\alpha_i$.

The proof for the asymptotic behaviours of the return probabilities is split up over the following sections. In Section 3 we calculate the asymptotics in the case when $\Psi(\bar{\theta}) > 0$,

$G_1'(\mathbf{r}_1) < \infty$ and $G_2'(\mathbf{r}_2) < \infty$ hold; see Theorem 3.1. In Section 4 we investigate the case when $\Psi(\bar{\theta}) = 0$, $G_1'(\mathbf{r}_1) < \infty$ and $G_2'(\mathbf{r}_2) < \infty$ hold; see Theorem 4.1. From the proof of this theorem we will see that even the case $\Psi(\bar{\theta}) = 0$, $G_1'(\zeta_1(\mathbf{r}_1)) < \infty$ and $G_2'(\zeta_2(\mathbf{r}_2)) < \infty$ is covered. In Section 5 we treat the remaining cases: Theorem 5.1 covers the case when $G_1(\mathbf{r}_1) < \infty$, $G_1'(\mathbf{r}_1) = \infty$ and $G_2'(\mathbf{r}_2) < \infty$ hold, while Corollary 5.2 answers the question for the asymptotic behaviour when $G_1'(\mathbf{r}_1) = \infty$ and $G_2'(\mathbf{r}_2) = \infty$. Finally, Theorem 5.3 covers the remaining case when $G_1(\mathbf{r}_1) = \infty$ or $G_2(\mathbf{r}_2) = \infty$.

3. THE ASYMPTOTIC BEHAVIOUR IN THE CASE $\Psi(\bar{\theta}) > 0$

Throughout this section we investigate the case $m = 2$ and assume that $\Psi(\bar{\theta}) > 0$ and $G_1(z)$ and $G_2(z)$ are differentiable at their radii of convergence. That is, the Green functions have an expansion as assumed in (2.2). Recall that the smallest singular term w.r.t. $\preceq$ in the expansion of $G_i(z)$, is denoted by $S_i(z) = (\mathbf{r}_i - z)^{q_i} \log^{k_i}(\mathbf{r}_i - z)$ with $d_i < q_i \le d_i + 1$. Let us remark that Darboux's method yields that the n-step return probabilities of the random walk on Γ_i governed by μ_i behave asymptotically like $\hat{C}_i \mathbf{r}_i^{-n\delta_i} n^{-\lambda_i} \log^{\kappa_i} n$, where λ_i and κ_i are given by (2.3). The aim of this section is to prove the following:

Theorem 3.1. *Assume that $G_1(z)$ and $G_2(z)$ are differentiable at $z = \mathbf{r}_1$, $z = \mathbf{r}_2$ respectively, and have an expansion as in (2.2). If $S_1(z) \preceq S_2(z)$ and $\Psi(\bar{\theta}) > 0$ then:*

$$\mu^{(n\delta)}(e) \sim \begin{cases} C_1 \cdot \mathbf{r}^{-n\delta} \cdot n^{-\lambda_1} \cdot \log^{\kappa_1}(n), & \text{if } \alpha_1 \ge \frac{\theta_1}{\theta_1+\theta_2}, \\ C_2 \cdot \mathbf{r}^{-n\delta} \cdot n^{-\lambda_2} \cdot \log^{\kappa_2}(n), & \text{if } \alpha_1 < \frac{\theta_1}{\theta_1+\theta_2}, \end{cases} \quad \text{for some constants } C_1, C_2 > 0.$$

In the following we may assume w.l.o.g. that $\theta = \bar{\theta} = \theta_1/\alpha_1$. Recall that $F(s_i|z) = F_i(s_i|\zeta_i(z))$ for all $s_i \in \operatorname{supp}(\mu_i)$. Then we rewrite (2.5) as follows:

$$\alpha_1 z = \zeta_1(z)\left(1 - \alpha_2 z \sum_{s_2 \in \operatorname{supp}(\mu_2)} \mu_2(s_2) F_2(s_2|\zeta_2(z))\right), \tag{3.1}$$

$$\alpha_2 z = \zeta_2(z)\left(1 - \alpha_1 z \sum_{s_1 \in \operatorname{supp}(\mu_1)} \mu_1(s_1) F_1(s_1|\zeta_1(z))\right). \tag{3.2}$$

Recall that $\zeta_1(\mathbf{r}) = \mathbf{r}_1$ and $\zeta_2(\mathbf{r}) \le \mathbf{r}_2$ with equality if and only if $\theta = \theta_1/\alpha_1 = \theta_2/\alpha_2$. We remark also that $\Psi(\bar{\theta}) > 0$ implies $G'(\mathbf{r}) < \infty$: since $\Phi'(\bar{\theta}) < \Phi(\bar{\theta})/\bar{\theta} = 1/\mathbf{r}$ we get by differentiating (2.7)

$$G'(\mathbf{r}) = \lim_{z \to \mathbf{r}} \frac{\Phi'(zG(z))\, G(z)}{1 - z\, \Phi'(zG(z))} = \frac{\Phi'(\bar{\theta})\, G(\mathbf{r})}{1 - \mathbf{r}\, \Phi'(\bar{\theta})} < \infty.$$

Furthermore, we define

$$D := \begin{cases} d_1, & \text{if } \bar{\theta} < \theta_2/\alpha_2, \\ \min\{d_1, d_2\}, & \text{if } \bar{\theta} = \theta_1/\alpha_1 = \theta_2/\alpha_2. \end{cases}$$

We denote by $S(z)$ the *main leading singular term*, which is given by

$$S(z) = \begin{cases} S_1(z), & \text{if } \bar{\theta} < \theta_2/\alpha_2, \\ \min\{S_1(z), S_2(z)\}, & \text{if } \bar{\theta} = \theta_2/\alpha_2. \end{cases}$$

Lemma 3.2. $0 < \zeta_1'(\mathbf{r}) < \infty$ *and* $0 < \zeta_2'(\mathbf{r}) < \infty$.

Proof. We prove the lemma only for $\zeta_1'(\mathbf{r})$. We write

$$H_2(z) := \alpha_2 z \sum_{s_2 \in \mathrm{supp}(\mu_2)} \mu_2(s_2) F_2(s_2|\zeta_2(z)).$$

Since $\zeta_1(\mathbf{r}) = \mathbf{r}_1$, we have $H_2(\mathbf{r}) < 1$; compare with the definition of $\zeta_1(z)$. Furthermore, the coefficient of z^n in $H_2(z)$ is just the probability for the random walk on Γ of starting at e, making the first step w.r.t. μ_2 and returning for the first time to e at time n. Thus, this probability is bounded from above by $\mu^{(n)}(e)$, and consequently $H_2'(\mathbf{r}) < G'(\mathbf{r}) < \infty$. Computing the derivative of $\zeta_1(z)$ in a neighbourhood of $z = \mathbf{r}$ gives

$$\zeta_1'(z) = \frac{\alpha_1(1 - H_2(z)) + \alpha_1 z H_2'(z)}{(1 - H_2(z))^2} > 0.$$

Finiteness of $\zeta_1'(\mathbf{r})$ follows now directly from the remarks above. ∎

The functions $F_i(s_i|z)$, where $i \in \{1, 2\}$ and $s_i \in \mathrm{supp}(\mu_i)$, are at least d_i-times differentiable at $z = \mathbf{r}_i$, since the same holds for $G_i(z)$ and we can compare the coefficients of z^n in the definitions of $F_i(s_i|z)$ and $G_i(z)$ as follows:

$$\mu_i^{(n)}(e_i) \geq \mu_i(s_i) \cdot \mathbb{P}\big[X_n^{(i)} = e_i, \forall m < n : X_m^{(i)} \neq e_i \,\big|\, X_0^{(i)} = s_i\big].$$

Thus, we can rewrite these functions in the form

$$F_i(s_i|z) = \sum_{n=0}^{d_i} f_n(s_i)(\mathbf{r}_i - z)^n + E^{(i)}(s_i|z) \tag{3.3}$$

with coefficients $f_n(s_i) \in \mathbb{R}$ and $E^{(i)}(s_i|z) = \mathbf{o}((\mathbf{r}_i - z)^{d_i})$. If $\zeta_2(\mathbf{r}) < \mathbf{r}_2$ then $F_2(s_2|z)$ is analytic at $z = \zeta_2(\mathbf{r})$ for all $s_2 \in \mathrm{supp}(\mu_2)$ and we can even write

$$F_2(s_2|z) = \sum_{n \geq 0} f_n(s_2)(\zeta_2(\mathbf{r}) - z)^n.$$

Now we can prove:

Lemma 3.3. *For $z \in \mathbb{C}$ in a neighbourhood of $\mathbf{r}_i$,*

$$\begin{aligned}&\sum_{s_i \in \mathrm{supp}(\mu_i)} \mu_i(s_i)\, z\, E^{(i)}(s_i|z)\\ &\quad = e^{(i)}_{(q_i,k_i)}(\mathbf{r}_i - z)^{q_i} \log^{k_i}(\mathbf{r}_i - z) + \sum_{(q,k)\in\widehat{\mathcal{T}}_i} e^{(i)}_{(q,k)}(\mathbf{r}_i - z)^q \log^k(\mathbf{r}_i - z) + \mathcal{O}((\mathbf{r}_i - z)^{d_i+2}),\end{aligned}$$

where $e^{(i)}_{(q_i,k_i)} \neq 0$ and $\widehat{\mathcal{T}}_i \subseteq \{(q,k) \in \mathbb{R}\times\mathbb{N}_0 \mid d_i < q \leq d_i+2, q > q_i$ or $(q = q_i \Rightarrow k < k_i)\}$ is finite.

Proof. Define

$$U_i(z) := \sum_{s_i \in \mathrm{supp}(\mu_i)} \mu_i(s_i)\, z\, F_i(s_i|z).$$

Observe that the expansions of $U_i(z)$ and $G_i(z)$ have the same leading singular term: indeed, both functions are d_i-times differentiable in a neighbourhood of $z = \mathbf{r}_i$ due to the well-known equation $G_i(z) = 1/(1 - U_i(z))$. Therefore, we have expansions

$$G_i(z) = \sum_{k=0}^{d_i} g_k^{(i)} (\mathbf{r}_i - z)^k + R_{G_i}(z) \quad \text{and} \quad U_i(z) = \sum_{k=0}^{d_i} u_k^{(i)} (\mathbf{r}_i - z)^k + R_{U_i}(z),$$

where $R_{G_i}(z) = \mathcal{O}_c(S_i(z))$ and $R_{U_i}(z) = \mathbf{o}((\mathbf{r}_i - z)^{d_i})$. Substituting these expansions into $G_i(z)(1 - U_i(z)) = 1$, and taking all polynomial terms to one side, we get

$$(1 - U_i(\mathbf{r}_i))\, R_{G_i}(z) - G_i(\mathbf{r}_i)\, R_{U_i(z)} = p(z) + \mathbf{o}((\mathbf{r}_i - z)^{d_i+1}),$$

where $p(z)$ is some polynomial. This equation implies that the right hand side is of order $\mathcal{O}((\mathbf{r}_i - z)^{d_i+1})$, that is, $R_{U_i(z)} = \mathcal{O}_c(S_i(z))$ and we can write

$$U_i(z) = \sum_{k=0}^{d_i} u_k^{(i)} (\mathbf{r}_i - z)^k + u_{(q_i,k_i)}^{(i)} S_i(z) + \widehat{R}_{U_i}(z) \quad \text{with } \widehat{R}_{U_i}(z) = \mathbf{o}(S_i(z)).$$

Plugging this expansion once again into $G_i(z)(1 - U_i(z)) = 1$, comparing error terms and iterating the last steps, together with substituting (3.3) in the definition of $U_i(z)$, yields the claim. ∎

The next goal is to show that $\zeta_1(z)$ and $\zeta_2(z)$ are D-times differentiable at $z = \mathbf{r}$.

Proposition 3.4. *There are real numbers $x_0, x_1, \ldots, x_D$ and $y_0, y_1, \ldots, y_D$ such that*

$$\zeta_1(z) = \sum_{k=0}^{D} x_k\, (\mathbf{r} - z)^k + X_D^{(1)}(z) \quad \textit{and} \quad \zeta_2(z) = \sum_{k=0}^{D} y_k\, (\mathbf{r} - z)^k + X_D^{(2)}(z),$$

where $X_D^{(1)}(z) = \mathbf{o}((\mathbf{r} - z)^D)$ and $X_D^{(2)}(z) = \mathbf{o}((\mathbf{r} - z)^D)$.

Proof. We prove the proposition by determining $x_0, x_1, \ldots, x_D$ and $y_0, y_1, \ldots, y_D$ inductively. By Lemma 3.2 and a well-known characterization of differentiability, we can rewrite $\zeta_1(z)$ and $\zeta_2(z)$ in the following way:

$$\begin{aligned} \zeta_1(z) &= \mathbf{r}_1 - \zeta_1'(\mathbf{r})\,(\mathbf{r} - z) + X_1^{(1)}(z), \text{where } X_1^{(1)}(z) = \mathbf{o}(\mathbf{r} - z), \\ \zeta_2(z) &= \zeta_2(\mathbf{r}) - \zeta_2'(\mathbf{r})\,(\mathbf{r} - z) + X_1^{(2)}(z), \text{where } X_1^{(2)}(z) = \mathbf{o}(\mathbf{r} - z). \end{aligned} \tag{3.4}$$

Thus, we have determined x_0, x_1 and y_0, y_1. Assume now that we can write for some $t < D$

$$\zeta_1(z) = \sum_{k=0}^{t} x_k\, (\mathbf{r} - z)^k + X_t^{(1)}(z) \text{ and } \zeta_2(z) = \sum_{k=0}^{t} y_k\, (\mathbf{r} - z)^k + X_t^{(2)}(z), \tag{3.5}$$

where $X_t^{(1)}(z) = \mathbf{o}((\mathbf{r} - z)^t)$ and $X_t^{(2)}(z) = \mathbf{o}((\mathbf{r} - z)^t)$. Recall from (3.3) that we have expansions of $F_1(s_1|z)$ and $F_2(s_2|z)$ of the form

$$\begin{aligned} F_1(s_1|z) &= \sum_{n=0}^{D} a_n(s_1)(\mathbf{r}_1 - z)^n + E^{(1)}(s_1|z) \text{ and} \\ F_2(s_2|z) &= \sum_{n=0}^{D} b_n(s_2)(\zeta_2(\mathbf{r}) - z)^n + E^{(2)}(s_2|z), \end{aligned} \tag{3.6}$$

where $E^{(i)}(s_i|z) = \mathbf{o}((\zeta_i(\mathbf{r}) - z)^D)$. In particular, if $\bar{\theta} < \theta_2/\alpha_2$ then $\zeta_2(\mathbf{r}) < \mathbf{r}_2$ and consequently we can even write $F_2(s_2|z) = \sum_{n\geq 0} b_n(s_2)(\zeta_2(\mathbf{r}) - z)^n$. Recall that the case $D = d_1 > d_2$ implies $\bar{\theta} < \theta_2/\alpha_2$. We now substitute the expansions (3.5) and (3.6) in Equations (3.1) and (3.2), yielding the following system:

$$
\begin{aligned}
\alpha_1 z &= \left(\sum_{k=0}^{t} x_k\,(\mathbf{r}-z)^k + X_t^{(1)}(z)\right)\left[1-\alpha_2(\mathbf{r}-(\mathbf{r}-z))\sum_{s_2\in\mathrm{supp}(\mu_2)}\mu_2(s_2)\cdot\right.\\
&\quad\left.\cdot\left[\sum_{n=0}^{D} b_n(s_2)\left(-\sum_{k=1}^{t} y_k\,(\mathbf{r}-z)^k - X_t^{(2)}(z)\right)^n + E^{(2)}(s_2|\zeta_2(z))\right]\right],\\
\alpha_2 z &= \left(\sum_{k=0}^{t} y_k\,(\mathbf{r}-z)^k + X_t^{(2)}(z)\right)\left[1-\alpha_1(\mathbf{r}-(\mathbf{r}-z))\sum_{s_1\in\mathrm{supp}(\mu_1)}\mu_1(s_1)\cdot\right.\\
&\quad\left.\cdot\left[\sum_{n=0}^{D} a_n(s_1)\left(-\sum_{k=1}^{t} x_k\,(\mathbf{r}-z)^k - X_t^{(1)}(z)\right)^n + E^{(1)}(s_1|\zeta_1(z))\right]\right].
\end{aligned}
\tag{3.7}
$$

Observe that $\sum_{s_i\in\mathrm{supp}(\mu_i)}\mu(s_i)zE^{(i)}(s_i|\zeta_i(z)) = \mathbf{o}((\zeta_i(\mathbf{r}) - \zeta_i(z))^D) = \mathbf{o}((\mathbf{r} - z)^D)$. We now bring all polynomial and higher order terms to the left hand side and get:

$$
\begin{aligned}
P_t^{(1)}(z) + \mathbf{o}((\mathbf{r}-z)^{t+1}) &= \left[1-\alpha_2\mathbf{r}\sum_{s_2\in\mathrm{supp}(\mu_2)}\mu_2(s_2)b_0(s_2)\right]X_t^{(1)}(z)\\
&\quad+\left[\alpha_2\mathbf{r}_1\mathbf{r}\sum_{s_2\in\mathrm{supp}(\mu_2)}\mu_2(s_2)b_1(s_2)\right]X_t^{(2)}(z),\\
P_t^{(2)}(z) + \mathbf{o}((\mathbf{r}-z)^{t+1}) &= \left[\alpha_1\zeta_2(\mathbf{r})\mathbf{r}\sum_{s_1\in\mathrm{supp}(\mu_1)}\mu_1(s_1)a_1(s_1)\right]X_t^{(1)}(z)\\
&\quad+\left[1-\alpha_1\mathbf{r}\sum_{s_1\in\mathrm{supp}(\mu_1)}\mu_1(s_1)a_0(s_1)\right]X_t^{(2)}(z),
\end{aligned}
\tag{3.8}
$$

where $P_t^{(1)}(z)$ and $P_t^{(2)}(z)$ are polynomials in the variable z. By assumption on $X_t^{(1)}(z)$ and $X_t^{(2)}(z)$, the right hand sides of (3.8) are of order $\mathbf{o}((\mathbf{r}-z)^t)$. Therefore, the left hand sides have to be of order $\mathcal{O}((\mathbf{r}-z)^{t+1})$, and consequently the right hand sides have to be also of order $\mathcal{O}((\mathbf{r}-z)^{t+1})$. It remains to show that $X_t^{(1)}(z) = \mathcal{O}((\mathbf{r}-z)^{t+1})$ and $X_t^{(2)}(z) = \mathcal{O}((\mathbf{r}-z)^{t+1})$. For this purpose, define the matrix $M = (m_{ij})_{1\leq i,j\leq 2}$ by

$$
\begin{aligned}
m_{11} &:= 1-\alpha_2\mathbf{r}\sum_{s_2\in\mathrm{supp}(\mu_2)}\mu_2(s_2)b_0(s_2),\\
m_{12} &:= \alpha_2\mathbf{r}_1\mathbf{r}\sum_{s_2\in\mathrm{supp}(\mu_2)}\mu_2(s_2)b_1(s_2),
\end{aligned}
$$

$$m_{21} := \alpha_1 \zeta_2(\mathbf{r})\mathbf{r} \sum_{s_1 \in \mathrm{supp}(\mu_1)} \mu_1(s_1) a_1(s_1),$$
$$m_{22} := 1 - \alpha_1 \mathbf{r} \sum_{s_1 \in \mathrm{supp}(\mu_1)} \mu_1(s_1) a_0(s_1).$$

Then the system (3.8) is equivalent to

$$M \cdot \begin{pmatrix} X_t^{(1)}(z) \\ X_t^{(2)}(z) \end{pmatrix} = \begin{pmatrix} Q_t^{(1)}(z) \\ Q_t^{(2)}(z) \end{pmatrix},$$

where $Q_t^{(1)}(z) = \mathcal{O}((\mathbf{r}-z)^{t+1})$ and $Q_t^{(2)}(z) = \mathcal{O}((\mathbf{r}-z)^{t+1})$. If the matrix M is invertible, then obviously $X_t^{(1)}(z) = \mathcal{O}((\mathbf{r}-z)^{t+1})$ and $X_t^{(2)}(z) = \mathcal{O}((\mathbf{r}-z)^{t+1})$. To this end, we now prove invertibility of M:

Lemma 3.5. $\det(M) \neq 0$.

Proof. We start with differentiating equations (3.1) and (3.2):

$$\alpha_1 = \left(-\alpha_2 \sum_{s_2 \in \mathrm{supp}(\mu_2)} \mu_2(s_2) F_2(s_2|\zeta_2(z)) - \alpha_2 z \sum_{s_2 \in \mathrm{supp}(\mu_2)} \mu_2(s_2) F_2'(s_2|\zeta_2(z))\zeta_2'(z) \right) \zeta_1(z)$$
$$+ \zeta_1'(z) \left(1 - \alpha_2 z \sum_{s_2 \in \mathrm{supp}(\mu_2)} \mu_2(s_2) F_2(s_2|\zeta_2(z)) \right),$$
$$\alpha_2 = \left(-\alpha_1 \sum_{s_1 \in \mathrm{supp}(\mu_1)} \mu_1(s_1) F_1(s_1|\zeta_1(z)) - \alpha_1 z \sum_{s_1 \in \mathrm{supp}(\mu_1)} \mu_1(s_1) F_1'(s_1|\zeta_1(z))\zeta_1'(z) \right) \zeta_2(z)$$
$$+ \zeta_2'(z) \left(1 - \alpha_1 z \sum_{s_1 \in \mathrm{supp}(\mu_1)} \mu_1(s_1) F_1(s_1|\zeta_1(z)) \right).$$

Observe that we have $a_0(s_1) = F_1(s_1|\mathbf{r}_1)$, $a_1(s_1) = -F_1'(s_1|\mathbf{r}_1)$, $b_0(s_2) = F_2(s_2|\zeta_2(\mathbf{r}))$ and $b_1(s_2) = -F_2'(s_2|\zeta_2(\mathbf{r}))$. Substituting these values in the above system and letting $z \to \mathbf{r}$ yields

$$\alpha_1 = \left(-\alpha_2 \sum_{s_2 \in \mathrm{supp}(\mu_2)} \mu_2(s_2) b_0(s_2) + \alpha_2 \mathbf{r} \sum_{s_2 \in \mathrm{supp}(\mu_2)} \mu_2(s_2) b_1(s_2) \zeta_2'(\mathbf{r}) \right) \mathbf{r}_1 + \zeta_1'(\mathbf{r}) m_{11},$$
$$\alpha_2 = \left(-\alpha_1 \sum_{s_1 \in \mathrm{supp}(\mu_1)} \mu_1(s_1) a_0(s_1) + \alpha_1 \mathbf{r} \sum_{s_1 \in \mathrm{supp}(\mu_1)} \mu_1(s_1) a_1(s_1) \zeta_1'(\mathbf{r}) \right) \zeta_2(\mathbf{r}) + \zeta_2'(\mathbf{r}) m_{22}.$$

Since $\zeta_1(\mathbf{r}), \zeta_2(\mathbf{r}) > 0$ and $a_1(s_1), b_1(s_2) < 0$ the last equations imply $m_{11}, m_{22} > 0$. We proceed with rewriting the last system:

$$\alpha_2 \mathbf{r}_1 \mathbf{r} \sum_{s_2 \in \mathrm{supp}(\mu_2)} \mu_2(s_2) b_1(s_2) \zeta_2'(\mathbf{r}) = A - \zeta_1'(\mathbf{r}) m_{11},$$
$$\alpha_1 \zeta_2(\mathbf{r}) \mathbf{r} \sum_{s_1 \in \mathrm{supp}(\mu_1)} \mu_1(s_1) a_1(s_1) \zeta_1'(\mathbf{r}) = B - \zeta_2'(\mathbf{r}) m_{22}, \tag{3.9}$$

where

$$A := \alpha_1 + \alpha_2 \mathbf{r}_1 \sum_{s_2 \in \mathrm{supp}(\mu_2)} \mu_2(s_2) b_0(s_2) \text{ and } B := \alpha_2 + \alpha_1 \zeta_2(\mathbf{r}) \sum_{s_1 \in \mathrm{supp}(\mu_1)} \mu_1(s_1) a_0(s_1).$$

Multiplying both equations in (3.9) yields the equation

$$\zeta_1'(\mathbf{r})\zeta_2'(\mathbf{r}) m_{12} m_{21} = AB - \zeta_1'(\mathbf{r}) m_{11} B - \zeta_2'(\mathbf{r}) m_{22} A + \zeta_1'(\mathbf{r})\zeta_2'(\mathbf{r}) m_{11} m_{22}.$$

Assume now that $\det(M) = 0$. Then we would get

$$\zeta_1'(\mathbf{r}) m_{11} B + \zeta_2'(\mathbf{r}) m_{22} A = AB,$$

or equivalently,

$$\zeta_2'(\mathbf{r}) = \frac{AB - \zeta_1'(\mathbf{r}) m_{11} B}{m_{22} A}. \tag{3.10}$$

Furthermore, (3.9) implies

$$\zeta_1'(\mathbf{r}) = \big(A - C\zeta_2'(\mathbf{r})\big)/m_{11},$$

where $C := \alpha_2 \mathbf{r}_1 \mathbf{r} \sum_{s_2 \in \mathrm{supp}(\mu_2)} \mu_2(s_2) b_1(s_2) < 0$. Substituting the last equation in (3.10) would lead to

$$\zeta_2'(\mathbf{r}) = \frac{BC}{m_{22} A} \zeta_2'(\mathbf{r}).$$

Observe now that $A, B, m_{22} > 0$ and $C < 0$. This yields a contradiction in the last equation, since $\zeta_2'(\mathbf{r}) > 0$. Thus, $\det(M) \neq 0$. ▪

The last lemma finishes the proof of Proposition 3.4. ▪

Recall the definition of the main leading singular term $S(z) = S_i(z) = (\mathbf{r}_i - z)^{q_i} \log^{k_i}(\mathbf{r}_i - z)$. The next aim is to show that at least one of the functions $X_D^{(1)}(z)$ and $X_D^{(2)}(z)$ has order $\mathcal{O}_c((\mathbf{r} - z)^{q_i} \log^{k_i}(\mathbf{r} - z))$. To this end, we look at the final step of the induction in the proof of Proposition 3.4. For $t = D$, the system (3.7) becomes

$$\left[1 - \alpha_2 \mathbf{r} \sum_{s_2 \in \mathrm{supp}(\mu_2)} \mu_2(s_2) b_0(s_2)\right] \cdot X_D^{(1)}(z) + \left[\alpha_2 \mathbf{r}_1 \mathbf{r} \sum_{s_2 \in \mathrm{supp}(\mu_2)} \mu_2(s_2) b_1(s_2)\right] \cdot X_D^{(2)}(z)$$
$$- \alpha_2 \mathbf{r}_1 \sum_{s_2 \in \mathrm{supp}(\mu_2)} \mu_2(s_2)\, z\, E^{(2)}(s_2 | \zeta_2(z)) = P_D^{(1)}(z) + \mathbf{o}((\mathbf{r} - z)^{D+1}),$$
$$\left[\alpha_1 \zeta_2(\mathbf{r}) \mathbf{r} \sum_{s_1 \in \mathrm{supp}(\mu_1)} \mu_1(s_1) a_1(s_1)\right] \cdot X_D^{(1)}(z) + \left[1 - \alpha_1 \mathbf{r} \sum_{s_1 \in \mathrm{supp}(\mu_1)} \mu_1(s_1) a_0(s_1)\right] \cdot X_D^{(2)}(z)$$
$$- \alpha_1 \zeta_2(\mathbf{r}) \sum_{s_1 \in \mathrm{supp}(\mu_1)} \mu_1(s_1)\, z\, E^{(1)}(s_1 | \zeta_1(z)) = P_D^{(2)}(z) + \mathbf{o}((\mathbf{r} - z)^{D+1}),$$

where $P_D^{(1)}(z)$ and $P_D^{(2)}(z)$ are polynomials in the variable z. By (3.4), we may conclude that $(\zeta_i(\mathbf{r}) - \zeta_i(z)) = \mathcal{O}_c(\mathbf{r} - z)$. Since $\zeta_i'(\mathbf{r}_i) < \infty$ by Lemma 3.2, we have for $1 < p \in \mathbb{R}$

$$(\zeta_i(\mathbf{r}) - \zeta_i(z))^p = \big(\zeta_i'(\mathbf{r}_i)(\mathbf{r} - z) + \mathbf{o}(\mathbf{r} - z)\big)^p = \zeta_i'(\mathbf{r}_i)^p\, (\mathbf{r} - z)^p\, (1 + \mathbf{o}(1))^p = \mathcal{O}_c((\mathbf{r} - z)^p)$$

and

$$\begin{aligned}\log(\zeta_i(\mathbf{r}) - \zeta_i(z)) &= \log\left(\zeta_i'(\mathbf{r}_i)(\mathbf{r} - z) + \mathbf{o}(\mathbf{r} - z)\right)\\ &= \log\left(\zeta_i'(\mathbf{r}_i)\right) + \log(\mathbf{r} - z) + \log(1 + \mathbf{o}(1))\\ &= \log\left(\zeta_i'(\mathbf{r}_i)\right) + \log(\mathbf{r} - z) + \mathbf{o}(1).\end{aligned}$$

We remark that $(1 + z)^p$ and $\log(1 + z)$ are analytic in a neighbourhood of $z = 0$. In the following we denote by $i \in \{1, 2\}$ the index such that $S(z) = S_i(z)$. Then, the computations above imply with Lemma 3.3 that

$$\sum_{s_i \in \text{supp}(\mu_i)} \mu(s_i)\, z\, E^{(i)}(s_i|\zeta_i(z)) = \mathcal{O}_c((\mathbf{r} - z)^{q_i} \log^{k_i}(\mathbf{r} - z)).$$

Since the matrix M from the proof of Proposition 3.4 is invertible, we can conclude analogously that we must have

$$X_D^{(1)}(z) = \mathcal{O}_c((\mathbf{r} - z)^{q_i} \log^{k_i}(\mathbf{r} - z)) \text{ and } X_D^{(2)}(z) = \mathcal{O}_c((\mathbf{r} - z)^{q_i} \log^{k_i}(\mathbf{r} - z)).$$

Thus, the leading singular term of $\zeta_i(z)$ has the same order as the leading singular term in the expansion of $G_i(z)$ if $S(z) = S_i(z)$. By (2.6), we can conclude that the leading singular term in the expansion of $G(z)$ at $z = \mathbf{r}$ has the same form as the leading singular term in the expansion of $G_i(z)$ at $z = \mathbf{r}_i$, namely $(\mathbf{r} - z)^{q_i} \log^{k_i}(\mathbf{r} - z)$.

Recall that we assumed throughout this section that $G_i(z)$ is exactly d_i-times differentiable at $z = \mathbf{r}_i$. For an application of *Darboux's method* we need in a first step the expansion of $G(z)$ in a neighbourhood of $z = \mathbf{r}$ up to terms of order $(\mathbf{r} - z)^{D+2}$, where $D = d_1$, if $\bar{\theta} < \theta_2/\alpha_2$, and $D = \min\{d_1, d_2\}$, if $\bar{\theta} = \theta_1/\alpha_1 = \theta_2/\alpha_2$. Thus, by (2.6), we have to extend the expansions of $\zeta_1(z)$ and $\zeta_2(z)$ up to terms of order $(\mathbf{r} - z)^{D+2}$. The next lemma ensures that there are only finitely many terms up to order $(\mathbf{r} - z)^{D+2}$ in these expansions.

Lemma 3.6. *For $i \in \{1, 2\}$, $\zeta_i(z)$ has an expansion of the form*

$$\sum_{k=0}^{D} x_k (\mathbf{r} - z)^k + \sum_{(q,k) \in \mathcal{T}} x_{(q,k)} (\mathbf{r} - z)^q \log^k(\mathbf{r} - z) + \mathbf{o}((\mathbf{r} - z)^{D+2}),$$

where $x_k, x_{(q,k)} \in \mathbb{R}$, $\mathcal{T}$ is a finite subset of $\widehat{\mathcal{T}} := \{(q, k) \in \mathbb{R} \times \mathbb{N}_0 \mid D < q \leq D + 2\}$. In particular, $(q_i, k_i) \in \mathcal{T}$ with $x_{(q_i,k_i)} \neq 0$, and $(q, k) \in \mathcal{T}$ implies $(q_i, k_i) \preceq (q, k)$.

Proof. Recall the expansion of $\sum_{s_i \in \text{supp}(\mu_i)} \mu_i(s_i)\, z\, E^{(i)}(s_i|z)$ from Lemma 3.3. Assume that $\zeta_i(z)$ has already an expansion of the form

$$\sum_{k=0}^{D} x_k (\mathbf{r} - z)^k + \sum_{(q,k) \in \mathcal{T}'} x_{(q,k)} (\mathbf{r} - z)^q \log^k(\mathbf{r} - z) + \mathbf{o}(\max \mathcal{T}'), \tag{3.11}$$

where $\mathcal{T}'$ is a finite subset of $\widehat{\mathcal{T}}$ and $\max \mathcal{T}' := \max_{\preceq}\{(\mathbf{r} - z)^q \log^k(\mathbf{r} - z) \mid (q, k) \in \mathcal{T}'\}$. In particular, $x_{(q_i,k_i)} \in \mathcal{T}'$ with $x_{(q_i,k_i)} \neq 0$. We proceed with expanding the next terms of $\zeta_i(z)$

analogously to the proof of Proposition 3.4. For this purpose, observe that for $p > 1$ we can rewrite $(\zeta_i(\mathbf{r}) - \zeta_i(z))^p$ as

$$(-x_1)^p\,(\mathbf{r}-z)^p \\ \times \left(1+\sum_{k=2}^{D}\frac{x_k}{x_1}(\mathbf{r}-z)^{k-1}+\sum_{(q,k)\in\mathcal{T}'}\frac{x_{(q,k)}}{x_1}(\mathbf{r}-z)^{q-1}\log^k(\mathbf{r}-z)+\mathbf{o}\left(\frac{\max\mathcal{T}'}{\mathbf{r}-z}\right)\right)^p \tag{3.12}$$

and $\log(\zeta_i(\mathbf{r}) - \zeta_i(z))$ as

$$C+\log(\mathbf{r}-z) \\ +\log\left(1+\sum_{k=2}^{D}\frac{x_k}{x_1}(\mathbf{r}-z)^{k-1}+\sum_{(q,k)\in\mathcal{T}'}\frac{x_{(q,k)}}{x_1}(\mathbf{r}-z)^{q-1}\log^k(\mathbf{r}-z)+\mathbf{o}\left(\frac{\max\mathcal{T}'}{\mathbf{r}-z}\right)\right). \tag{3.13}$$

Note that $(1+z)^p$ with $p > 1$ and $\log(1+z)$ are analytic in a neighbourhood of $z = 0$. We substitute (3.11), (3.12) and (3.13) in Equations (3.1) and (3.2) and compare again the error terms (we will repeat this procedure in each of the following steps). Therefore, if $\max\mathcal{T}' = (\mathbf{r}-z)^{\hat{q}}\log^{\hat{k}}(\mathbf{r}-z)$ then the next possible terms up to order $(\mathbf{r}-z)^{\hat{q}}$ in the expansion may only be

$$(\mathbf{r}-z)^{\hat{q}}\log^{\hat{k}-1}(\mathbf{r}-z),(\mathbf{r}-z)^{\hat{q}}\log^{\hat{k}-2}(\mathbf{r}-z),\dots,(\mathbf{r}-z)^{\hat{q}}.$$

Analogously to the proof of Proposition 3.4 we determine step by step the corresponding coefficients of these terms. The next term in the expansion of $\zeta_i(z)$ has now the form $(\mathbf{r}-z)^{\check{q}}\log^{\check{k}}(\mathbf{r}-z)$, where $\check{q} > \hat{q}$ is a sum of elements from the finite set

$$\{1, q, q-1 \mid (q,\cdot)\in\mathcal{T}_1\cup\mathcal{T}_2\}$$

with $\mathcal{T}_i$ given as in (2.2). The value of $\check{q}$ is minimal such that $\check{q} > \hat{q}$. Due to (3.12) and (3.13) there is obviously a maximal $\check{k}\in\mathbb{N}_0$ such that $(\mathbf{r}-z)^{\check{q}}\log^{\check{k}}(\mathbf{r}-z)$ may be a non-vanishing next term in the expansion of $\zeta_i(z)$. Thus, we may iterate the last few steps again. Since there are only finitely many possible values for q such that a term of the form $(\mathbf{r}-z)^q\log^k(\mathbf{r}-z)$ may appear in the expansion up to order $(\mathbf{r}-z)^{D+2}$, we have shown that there are only finitely many terms up to order $(\mathbf{r}-z)^{D+2}$ in the expansion of $\zeta_i(z)$. ∎

With the last lemma we are now able to prove Theorem 3.1:

Proof of Theorem 3.1. We start by expanding $\zeta_1(z)$ and $\zeta_2(z)$ as in Proposition 3.4. If $\alpha_1 > \theta_1/(\theta_1+\theta_2)$ then $\bar{\theta} = \theta_1/\alpha_1 < \theta_2/\alpha_2$ and $\zeta_1(\mathbf{r}) = \mathbf{r}_1$, $\zeta_2(\mathbf{r}) < \mathbf{r}_2$, and consequently the leading singular term in the expansion of $\zeta_1(z)$ (and $\zeta_2(z)$) is then given by the term $S_1(z) = (\mathbf{r}-z)^{q_1}\log^{k_1}(\mathbf{r}-z)$. Analogously, if we have $\bar{\theta} = \theta_2/\alpha_2 < \theta_1/\alpha_1$, then $\zeta_2(\mathbf{r}) = \mathbf{r}_2$ and $\zeta_1(\mathbf{r}) < \mathbf{r}_1$, and the leading singular term is then $S_2(z) = (\mathbf{r}-z)^{q_2}\log^{k_2}(\mathbf{r}-z)$. If $\alpha_1 = \theta_1/(\theta_1+\theta_2)$ then $\bar{\theta} = \theta_1/\alpha_1 = \theta_2/\alpha_2$, $\zeta_1(\mathbf{r}) = \mathbf{r}_1$, $\zeta_2(\mathbf{r}) = \mathbf{r}_2$, and the leading singular term in the expansions of $\zeta_1(z)$ and $\zeta_2(z)$ is $S_j(z) = (\mathbf{r}-z)^{q_j}\log^{k_j}(\mathbf{r}-z)$, where $j = 1$, if $S_1(z) \preceq S_2(z)$, and $j = 2$, if $S_2(z) \prec S_1(z)$. For the rest of the proof, we denote by $i\in\{1,2\}$

the index such that $S(z) = S_i(z)$. Therefore, the expansion of the common leading singular term of $\zeta_1(z)$ and $\zeta_2(z)$, namely $S_i(z)$, in a neighbourhood of 0 has coefficients of asymptotic order proportional to $\mathbf{r}^{-n} n^{-\lambda_i} \log^{\kappa_i} n$.

We will use the technique which is called *Darboux's method*: recall that the *Riemann-Lebesgue-Lemma* states that if a function $H(z) = \sum_{n\geq 0} h_n z^n$ has radius of convergence $\mathbf{r}_H$ and if H is k-times continuously differentiable on its circle of convergence, then $h_n \mathbf{r}_H^n n^k \to 0$ as $n \to \infty$. Thus, one identifies all singularities on the circle of convergence and subtracts parts of the expansion near them such that the remaining part is sufficiently often differentiable on the circle. The asymptotics of the coefficients arise then from the main leading singular terms. We refer to Olver [17, Chap. 8, §9.2] for more details.

Lemma 3.6 assures that we have a singular expansion of $\zeta_1(z)$ up to terms of order $\lceil \lambda_i \rceil = \lceil q_i \rceil + 1 = D + 2$, which allows us to apply Darboux's method: we get the asymptotic behaviour of $\mu^{(n\delta)}(e)$ by plugging $\zeta_1(z)$ into Equation (2.6). Thus, the leading singular term in the expansion of $G(z)$ in a neighbourhood of $z = \mathbf{r}$ is the same as the one of $\zeta_1(z)$, namely $(\mathbf{r} - z)^{q_i} \log^{k_i}(\mathbf{r} - z)$. We have to show that the expansion of $G(z)$ at every singular point on the disc of convergence has the same form. The singularities are exactly the points $\mathbf{r}\exp(\mathrm{i}2\pi j/\delta)$ with $0 \leq j < \delta - 1$; see e.g. [22, Theorem 9.4]. Writing $z = \lambda \mathbf{r} \omega_j$, where $\omega_j = \exp(\mathrm{i}2\pi j/\delta)$ and $\lambda \in \mathbb{C}$ with $|\lambda| < 1$,

$$G(z) = G(\lambda \mathbf{r}\omega_j) = \sum_{n\geq 0} \mu^{(n\delta)}(e)(\lambda \mathbf{r}\omega_j)^{n\delta} = \sum_{n\geq 0} \mu^{(n\delta)}(e)(\lambda \mathbf{r})^{n\delta} = G(\lambda \mathbf{r}) = G(z/\omega_j).$$

Thus, for every $j \in \{0, 1, \dots, \delta - 1\}$, we have expansions of $G(z)$ in a neighbourhood of $z = \mathbf{r}\omega_j$ given by

$$G(z) = \sum_{k=0}^{D} g_k(\mathbf{r} - z/\omega_j)^k + \sum_{(q,k)\in\widehat{\mathcal{T}_i}} g_{(q,k)}(\mathbf{r} - z/\omega_j)^q \log^k(\mathbf{r} - z/\omega_j) + \mathcal{O}\left(\left(\mathbf{r}\omega_j - z\right)^{D+2}\right),$$

where $\widehat{\mathcal{T}_i}$ is a finite subset of $\{(q,k) \in \mathbb{R}\times\mathbb{N} \mid D < q \leq D+2, q > q_i \vee (q = q_i \Rightarrow k < k_i)\}$, $g_{(q_i,k_i)} \in \widehat{\mathcal{T}_i}$ with $g_{(q_i,k_i)} \neq 0$ and $(q,k) \in \widehat{\mathcal{T}_i}$ implies $(q_i, k_i) \preceq (q,k)$. Therefore, the difference

$$G(z) - \sum_{j=0}^{\delta-1} \sum_{(q,k)\in\widehat{\mathcal{T}_i}} g_{(q,k)}(\mathbf{r} - z/\omega_j)^q \log^k(\mathbf{r} - z/\omega_j)$$

is $(D+2)$-times differentiable on the circle of convergence. Observe now that the coefficients of the expansion of $(\mathbf{r} - z/\omega_j)^{q_i} \log^{k_i}(\mathbf{r} - z/\omega_j)$ in a neighbourhood of 0 behave asymptotically like $C\,(\mathbf{r}\omega_j)^{-n}\, n^{-\lambda_i}\, \log^{\kappa_i}(n)$. We can drop higher order terms in the above difference because the corresponding coefficients have higher asymptotic order. Since $G(z) = \sum_{n\geq 0} \mu^{(n)} z^n$, we can conclude that

$$\mu^{(n)}(e) \sim \sum_{j=0}^{\delta-1} C\, n^{-\lambda_i}\, \log^{\kappa_i}(n)\, \mathbf{r}^{-n}\, \omega_j^{-n}.$$

Observe that $\sum_{j=0}^{\delta-1} \omega_j^{-n} = \delta$ if δ divides n, and this sum is zero otherwise.

We note once again that the asymptotic behaviour of the coefficients in the expansion of the function $(\mathbf{r} - z)^{q_i} \log^{k_i}(\mathbf{r} - z)$ near 0 are well-known; see e.g. Flajolet and Sedgewick [7]. ▪

Let us remark that the reasoning in the above proof shows analogously the asymptotic behaviour $\mu_i^{(n)}(e_i) \sim \hat{C}_i\, \mathbf{r}_i^{-n}\, n^{-\lambda_i} \log^{\kappa_i} n$. That is, in the presented case of $\Psi(\bar{\theta}) > 0$, $G_1'(\mathbf{r}_1) < \infty$ and $G_2'(\mathbf{r}_2) < \infty$ the asymptotics are directly inherited from the asymptotics of the random walk on Γ_i governed by μ_i.

4. THE CASE $\Psi(\bar{\theta}) = 0$

We now consider the case $\Gamma = \Gamma_1 * \Gamma_2$ and assume that $\Psi(\bar{\theta}) = 0$, $G_1'(\mathbf{r}_1) < \infty$ and $G_2'(\zeta_2(\mathbf{r})) < \infty$ hold. W.l.o.g. we may also assume $\theta = \bar{\theta} = \theta_1/\alpha_1$. The aim of this section is to prove the following:

Theorem 4.1. *Assume that $G_1'(\mathbf{r}_1) < \infty$ and $G_2'(\zeta_2(\mathbf{r}_2)) < \infty$. If $\Psi(\bar{\theta}) = 0$ then*

$$\mu^{(n\delta)}(e) \sim C \cdot \mathbf{r}^{-n\delta} \cdot n^{-3/2}.$$

In the following we will derive expansions of $\zeta_i(z)$ and $G(z)$ in a neighbourhood of $z = \mathbf{r}$ in order to prove Theorem 4.1. Recall from (2.8) that $\Psi(\bar{\theta}) = 0$ implies

$$\Phi'(\bar{\theta}) = \frac{\Phi(\bar{\theta})}{\bar{\theta}} = \frac{\Phi(\theta)}{\theta} = \frac{\Phi(\mathbf{r}\, G(\mathbf{r}))}{\mathbf{r}\, G(\mathbf{r})} = \frac{G(\mathbf{r})}{\mathbf{r}\, G(\mathbf{r})} = \frac{1}{\mathbf{r}}.$$

Differentiating (2.7) yields

$$G'(z) = \frac{G(z)\Phi'(zG(z))}{1 - z\Phi'(zG(z))}. \tag{4.1}$$

Therefore, $G'(\mathbf{r}) = \infty$, and consequently we have to proceed differently from the previous section in order to find the expansion of $G(z)$. First, we show positivity of $\Phi''(\bar{\theta})$ in the present setting:

Lemma 4.2. *Assume that $G_1'(\mathbf{r}_1) < \infty$ and $G_2'(\zeta_2(\mathbf{r}_2))$. If $\Psi(\bar{\theta}) = 0$ then $\Phi''(\bar{\theta}) > 0$.*

Proof. Differentiating (2.9) twice yields

$$\Phi''(\bar{\theta}) = \alpha_1^2 \Phi_1''(\alpha_1\bar{\theta}) + \alpha_2^2 \Phi_2''(\alpha_2\bar{\theta}). \tag{4.2}$$

Since $\Phi_1(t)$ and $\Phi_2(t)$ are strictly convex for $t \in [0, \theta_1)$ and $t \in [0, \theta_2)$ respectively, we get $\Phi''(\bar{\theta}) > 0$ whenever $\theta_1/\alpha_1 \neq \theta_2/\alpha_2$: if $\bar{\theta} = \theta_1/\alpha_1 < \theta_2/\alpha_2$ then $\alpha_2\bar{\theta} < \theta_2$, that is, $\Phi_2''(\alpha_2\bar{\theta}) > 0$.

We now consider the case $\theta_1/\alpha_1 = \theta_2/\alpha_2$, that is, $\zeta_2(\mathbf{r}) = \mathbf{r}_2$. Assume now $\Phi''(\bar{\theta}) = 0$. Then $\Phi_1''(\theta_1) = \lim_{t\to\theta_1-} \Phi_1''(t) = 0$ and $\Phi_2''(\theta_2) = \lim_{t\to\theta_2-} \Phi_2''(t) = 0$ must hold. For $i \in \{1, 2\}$, differentiating (2.7) yields

$$G_i'(\mathbf{r}_i) = \lim_{z\to\mathbf{r}_i} \frac{G_i(z)\Phi_i'(zG_i(z))}{1 - z\Phi_i'(zG_i(z))},$$

or equivalently

$$\Phi_i'(\theta_i) = \lim_{z\to\mathbf{r}_i} \frac{G_i'(z)}{zG_i'(z) + G_i(z)} = \frac{G_i'(\mathbf{r}_i)}{\mathbf{r}_i G_i'(\mathbf{r}_i) + G_i(\mathbf{r}_i)} < \infty.$$

In particular, we have $\Phi_i'(\theta_i) < 1/\mathbf{r}_i$ since $G_i'(\mathbf{r}_i) < \infty$ by assumption. If $\Phi_i''(\theta_i) = 0$, differentiating (2.7) twice yields

$$G_i''(\mathbf{r}_i) = \lim_{z\to\mathbf{r}_i} \frac{\Phi_i''(zG_i(z))\big(G_i(z)+zG_i'(z)\big)^2 + 2\Phi_i'(zG_i(z))G_i(z)}{1-z\Phi_i'(zG_i(z))} = \frac{2\Phi_i'(\theta_i)G_i(\mathbf{r}_i)}{1-\mathbf{r}_i\Phi_i'(\theta_i)} < \infty.$$

Define the first return generating function as

$$U_i(z) := \sum_{n\geq 1} \mathbb{P}\big[X_n^{(i)} = e_i, \forall 1 \leq m \leq n-1 : X_m^{(i)} \neq e_i \mid X_0^{(i)} = e_i\big]\, z^n,$$

which satisfies the well-known equation $G_i(z) = 1/(1-U_i(z))$ and is strictly convex. $G_i''(\mathbf{r}_i) < \infty$ implies obviously $U_i''(\mathbf{r}_i) < \infty$. Therefore, we can compute $\Phi_i''(\theta_i)$ as

$$\Phi_i''(\theta_i) = \lim_{z\to\mathbf{r}_i} \frac{G_i(z)^3 U_i''(z)}{(G_i(z)+zG_i'(z))^3} = \frac{G_i(\mathbf{r}_i)^3 U_i''(\mathbf{r}_i)}{(G_i(\mathbf{r}_i)+\mathbf{r}_i G_i'(\mathbf{r}_i))^3} > 0,$$

a contradiction, and consequently $\Phi''(\bar{\theta}) > 0$ due to (4.2). ∎

We proceed with expanding $G(z)$ nearby $z = \mathbf{r}$.

Proposition 4.3. *Assume that $\Phi''(\bar{\theta}) < \infty$, $\Psi(\bar{\theta}) = 0$, $G_1'(\mathbf{r}_1) < \infty$ and $G_2'(\zeta_2(\mathbf{r})) < \infty$ hold. Then we can expand $G(z)$ in a neighbourhood of $z = \mathbf{r}$ as follows:*

$$G(z) = g_0 + g_1\sqrt{\mathbf{r}-z} + \mathbf{o}(\sqrt{\mathbf{r}-z}),$$

where $g_0, g_1 \in \mathbb{R}$ with $g_1 \neq 0$.

Proof. Consider the auxiliary function $H(z) := (G(z) - G(\mathbf{r}))^2$, and its first derivative $H'(z) = 2G'(z)(G(z) - G(\mathbf{r}))$. Using Equation (4.1), we get

$$H'(z) = 2\frac{G(z)\Phi'(zG(z))}{1-z\Phi'(zG(z))}(G(z) - G(\mathbf{r})).$$

The next aim is to show differentiability of $H(z)$ at $z = \mathbf{r}$. For this purpose, we want to show finiteness of the following limit:

$$\lim_{z\to\mathbf{r}} H'(z) = \lim_{z\to\mathbf{r}} 2G(z)\Phi'(zG(z))\frac{G(z)-G(\mathbf{r})}{1-z\Phi'(zG(z))}.$$

Since $2G(z)\Phi'(zG(z))$ tends to $A := 2G(\mathbf{r})/\mathbf{r} < \infty$, we just look at the following limit:

$$\begin{aligned}\lim_{z\to\mathbf{r}} \frac{G(z)-G(\mathbf{r})}{1-z\Phi'(zG(z))} &= \lim_{z\to\mathbf{r}} \frac{\Phi(zG(z)) - G(\mathbf{r})}{1-z\Phi'(zG(z))}\\ &= \lim_{z\to\mathbf{r}} \frac{\Phi'(zG(z))(G(z)+zG'(z))}{-\Phi'(zG(z)) - z\Phi''(zG(z))(G(z)+zG'(z))}. \qquad (4.3)\end{aligned}$$

In the last equation we applied De L'Hôpital's rule. We now write $\mathcal{G}(z) := G(z) + zG'(z)$, which tends to infinity for $z \to \mathbf{r}$. Recall that $\bar{\theta} = \theta = \mathbf{r}G(\mathbf{r})$ if $\Psi(\bar{\theta}) = 0$. Therefore, Equation (4.3) yields

$$H'(\mathbf{r}) = \lim_{z\to\mathbf{r}} \frac{A\Phi'(\theta)\mathcal{G}(z)}{-\Phi'(\theta) - \mathbf{r}\Phi''(\theta)\mathcal{G}(z)} = \lim_{x\to\infty} \frac{A\Phi'(\theta)x}{-\Phi'(\theta) - \mathbf{r}\Phi''(\theta)x} = \frac{A}{-\mathbf{r}^2\Phi''(\theta)} \in (-\infty, 0).$$

Thus,

$$\lim_{z\to\mathbf{r}} \frac{G(\mathbf{r})-G(z)}{\sqrt{\mathbf{r}-z}} = \lim_{z\to\mathbf{r}} \sqrt{\frac{(G(z)-G(\mathbf{r}))^2}{\mathbf{r}-z}} = \sqrt{-H'(\mathbf{r})} \in (0,\infty)$$

leads to the proposed expansion, namely

$$G(z) = G(\mathbf{r}) - \sqrt{-H'(\mathbf{r})}\sqrt{\mathbf{r}-z} + \mathbf{o}(\sqrt{\mathbf{r}-z}),$$

where $\sqrt{-H'(\mathbf{r})} \neq 0$. ∎

The next lemma shows that also $\zeta_1(z)$ and $\zeta_2(z)$ have the same expansion type:

Lemma 4.4. *Assume $\Phi''(\bar{\theta}) < \infty$. If $\Psi(\bar{\theta}) = 0$, $G_1'(\mathbf{r}_1) < \infty$ and $G_2'(\zeta_2(\mathbf{r})) < \infty$ we can expand $\zeta_1(z)$ and $\zeta_2(z)$ in a neighbourhood of $z = \mathbf{r}$ in the following way:*

$$\zeta_1(z) = \mathbf{r}_1 + a_0\sqrt{\mathbf{r}-z} + \mathbf{o}(\sqrt{\mathbf{r}-z}), \quad \zeta_2(z) = \zeta_2(\mathbf{r}) + b_0\sqrt{\mathbf{r}-z} + \mathbf{o}(\sqrt{\mathbf{r}-z}),$$

where $a_0, b_0 \in \mathbb{R}\setminus\{0\}$.

Proof. Obviously, we can write

$$\zeta_1(z) = \mathbf{r}_1 + X_1(z), \quad \zeta_2(z) = \zeta_2(\mathbf{r}) + X_2(z), \tag{4.4}$$

where $X_1(\mathbf{r}) = X_2(\mathbf{r}) = 0$. Moreover, for $i \in \{1,2\}$,

$$G_i(\zeta_i(z)) = G_i(\zeta_i(\mathbf{r})) - G_i'(\zeta_i(\mathbf{r}))(-X_i(z)) + \mathbf{o}(X_i(z)). \tag{4.5}$$

Substituting (4.4) and (4.5) in (2.6) yields the claim when comparing all error terms. ∎

Now we can show that $\Phi''(\bar{\theta}) < \infty$ holds in the present setting:

Lemma 4.5. *Assume $G_1'(\mathbf{r}_1) < \infty$ and $G_1(\zeta_2(\mathbf{r})) < \infty$. If $\Psi(\bar{\theta}) = 0$ then $\Phi''(\bar{\theta}) < \infty$.*

Proof. Assume now that $\Phi''(\bar{\theta}) = \infty$. We rewrite $\zeta_1(z)$ and $\zeta_2(z)$ as

$$\zeta_1(z) = \mathbf{r}_1 + X_1(z), \quad \text{and} \quad \zeta_2(z) = \zeta_2(\mathbf{r}) + X_2(z), \tag{4.6}$$

with $X_1(\mathbf{r}) = X_2(\mathbf{r}) = 0$. More precisely, if $\Phi''(\bar{\theta}) = \infty$, then the reasoning in Proposition 4.3 yields $H'(\mathbf{r}) = 0$, and consequently $X_1(z), X_2(z) = \mathbf{o}(\sqrt{\mathbf{r}-z})$. Furthermore, $X_1(z), X_2(z) \neq \mathcal{O}((\mathbf{r}-z))$, because otherwise $\zeta_1'(\mathbf{r}), \zeta_2'(\mathbf{r}) < \infty$ together with (2.6) would lead to a contradiction with $G'(\mathbf{r}) = \infty$. For $i \in \{1,2\}$ and $s_i \in \mathrm{supp}(\mu_i)$, we write in the following $F_i(s_i|z) = \sum_{n\geq 1} f_n^{(i)}(s_i)z^n$ with suitable coefficients $f_n^{(i)}(s_i) \in \mathbb{R}$. Our next aim is to find real numbers $C_1^{(i)}$ and $C_2^{(i)}$ such that

$$C_1^{(i)}X_1(z) + C_2^{(i)}X_2(z) + \mathbf{o}(\mathbf{r}-z) = \mathrm{LP}_i, \tag{4.7}$$

where LP_i is a linear polynomial. For this purpose, we rewrite Equations (3.1) and (3.2) with the help of (4.6). In the following denote by j the element of $\{1,2\}$ which is different from i. We get:

$$\left(1-\alpha_j(\mathbf{r}-(\mathbf{r}-z))\sum_{s_j\in\mathrm{supp}(\mu_j)}\mu_j(s_j)\sum_{n\geq 1}f_n^{(j)}(s_j)(\zeta_j(\mathbf{r})+X_j(z))^n\right)(\zeta_i(\mathbf{r})+X_i(z))=\alpha_i z. \tag{4.8}$$

The coefficients $C_1^{(i)}$ and $C_2^{(i)}$ of $X_1(z)$ and $X_2(z)$ respectively, are

$$\begin{aligned}
C_1^{(1)} &:= 1-\alpha_2\mathbf{r}\sum_{s_2\in\mathrm{supp}(\mu_2)}\mu_2(s_2)\sum_{n\geq 1}f_n^{(2)}(s_2)\,\zeta_2(\mathbf{r})^n\\
&= 1-\alpha_2\mathbf{r}\sum_{s_2\in\mathrm{supp}(\mu_2)}\mu_2(s_2)F_2(s_2|\zeta_2(\mathbf{r})),\\
C_2^{(1)} &:= -\alpha_2\mathbf{r}_1\mathbf{r}\sum_{s_2\in\mathrm{supp}(\mu_2)}\mu_2(s_2)\sum_{n\geq 1}f_n^{(2)}(s_2)n\,\zeta_2(\mathbf{r})^{n-1}\\
&= -\alpha_2\mathbf{r}_1\mathbf{r}\sum_{s_2\in\mathrm{supp}(\mu_2)}\mu_2(s_2)F_2'(s_2|\zeta_2(\mathbf{r})),\\
C_1^{(2)} &:= -\alpha_1\zeta_2(\mathbf{r})\mathbf{r}\sum_{s_1\in\mathrm{supp}(\mu_1)}\mu_1(s_1)\sum_{n\geq 1}f_n^{(1)}(s_1)n\mathbf{r}_1^{n-1}\\
&= -\alpha_1\zeta_2(\mathbf{r})\mathbf{r}\sum_{s_1\in\mathrm{supp}(\mu_1)}\mu_1(s_1)F_1'(s_1|\mathbf{r}_1),\\
C_2^{(2)} &:= 1-\alpha_1\mathbf{r}\sum_{s_1\in\mathrm{supp}(\mu_1)}\mu_1(s_1)\sum_{n\geq 1}f_n^{(1)}(s_1)\mathbf{r}_1^n\\
&= 1-\alpha_1\mathbf{r}\sum_{s_1\in\mathrm{supp}(\mu_1)}\mu_1(s_1)F_1(s_1|\mathbf{r}_1).
\end{aligned}$$

For $i=1$, the linear polynomial term on the left hand side of (4.8) is

$$\mathbf{r}_1\left(1-\alpha_2 z\sum_{s_2\in\mathrm{supp}(\mu_2)}\mu_2(s_2)F_2(s_2|\zeta_2(\mathbf{r}))\right),$$

while on the right hand side it is $\alpha_1 z$. For $i=2$, we have on the left hand side of (4.8)

$$\zeta_2(\mathbf{r})\left(1-\alpha_1 z\sum_{s_1\in\mathrm{supp}(\mu_1)}\mu_1(s_1)F_1(s_1|\mathbf{r}_1)\right),$$

and on the right hand side $\alpha_2 z$. Therefore, (4.7) holds with

$$\mathrm{LP}_1 := \alpha_1 z-\mathbf{r}_1\left(1-\alpha_2 z\sum_{s_2\in\mathrm{supp}(\mu_2)}\mu_2(s_2)F_2(s_2|\zeta_2(\mathbf{r}))\right)\text{ and}$$

$$\mathrm{LP}_2 := \alpha_2 z-\zeta_2(\mathbf{r})\left(1-\alpha_1 z\sum_{s_1\in\mathrm{supp}(\mu_1)}\mu_1(s_1)F_1(s_1|\mathbf{r}_1)\right).$$

The coefficients $C_1^{(i)}, C_2^{(i)}$ satisfy

$$C_1^{(1)}C_2^{(2)} - C_1^{(2)}C_2^{(1)} = 0. \tag{4.9}$$

Indeed, assume that $C_1^{(1)}C_2^{(2)} - C_1^{(2)}C_2^{(1)} \neq 0$. Then the following linear system

$$\begin{aligned} C_1^{(1)}X_1(z) + C_2^{(1)}X_2(z) + \mathbf{o}(\mathbf{r}-z) &= \mathrm{LP}_1, \\ C_1^{(2)}X_1(z) + C_2^{(2)}X_2(z) + \mathbf{o}(\mathbf{r}-z) &= \mathrm{LP}_2 \end{aligned}$$

has a unique solution for $X_1(z)$ and $X_2(z)$, but this means that both of them are of order $\mathcal{O}(\mathbf{r}-z)$, a contradiction to (4.6), where $X_1(z), X_2(z) \neq \mathcal{O}(\mathbf{r}-z)$.

Evaluating Equation (4.8) with $i=2$ at $z=\mathbf{r}$ gives $C_2^{(2)} > 0$. Equation (4.9) yields

$$\mathrm{LP}_1 - \frac{C_2^{(1)}}{C_2^{(2)}}\,\mathrm{LP}_2 = 0. \tag{4.10}$$

Evaluating the last equation at $z=0$ yields

$$-\mathbf{r}_1 + \frac{C_2^{(1)}}{C_2^{(2)}} \cdot \zeta_2(\mathbf{r}) = 0. \tag{4.11}$$

Since $C_2^{(1)} < 0$, Equation (4.11) gives us a contradiction, therefore $\Phi''(\bar{\theta}) = \infty$ cannot hold when $\Psi(\bar{\theta}) = 0$. ∎

We now proceed analogously to the previous section: we substitute the expansion of the last lemma in Equations (3.1) and (3.2) and determine step by step the next terms in the expansions of $\zeta_1(z)$ and $\zeta_2(z)$. The next lemma shows that we get only a finite number of terms up to order $(\mathbf{r}-z)^2$:

Lemma 4.6. *Let $i \in \{1,2\}$. If $\Psi(\bar{\theta}) = 0$, we can expand $\zeta_i(z)$ in a neighbourhood of $z=\mathbf{r}$ in the following way:*

$$\zeta_i(z) = \zeta_i(r) + c_0\sqrt{\mathbf{r}-z} + \sum_{(q,k)\in\mathcal{T}} c_{(q,k)}(\mathbf{r}-z)^q \log^k(\mathbf{r}-z) + \mathcal{O}((\mathbf{r}-z)^2),$$

where $\mathcal{T}$ is a finite subset of $\widehat{\mathcal{T}} := \{(q,k) \in \mathbb{R}\times\mathbb{N}_0 \mid 1/2 < q \leq 2\}$ and $c_0, c_{(q,k)} \in \mathbb{R}$ with $c_0 \neq 0$.

Proof. We start by plugging $\zeta_i(z) = \zeta_i(\mathbf{r}) + c_0\sqrt{\mathbf{r}-z} + X_0^{(i)}(z)$ with $X_0^{(i)}(z) = \mathbf{o}(\sqrt{r-z})$ into Equations (3.1) and (3.2) and determine step by step the next terms inductively analogously to the proof of Lemma 3.6. Assume now that $\zeta_i(z)$ has an expansion of the form

$$\zeta_i(\mathbf{r}) + c_0\sqrt{\mathbf{r}-z} + \sum_{(q,k)\in\mathcal{T}'} c_{(q,k)}(\mathbf{r}-z)^q \log^k(\mathbf{r}-z) + \mathbf{o}(\max\mathcal{T}'),$$

where $\mathcal{T}'$ with $\mathcal{T}' \subseteq \widehat{\mathcal{T}}$ finite. For $p>1$, $(\zeta_i(\mathbf{r}) - \zeta_i(z))^p$ can be rewritten as

$$(-c_0)^p\,(\mathbf{r}-z)^{p/2}\left(1 + \sum_{(q,k)\in\mathcal{T}'} \frac{c_{(q,k)}}{c_0}(\mathbf{r}-z)^{q-1/2}\log^k(\mathbf{r}-z) + \mathbf{o}\left(\frac{\max\mathcal{T}'}{\sqrt{\mathbf{r}-z}}\right)\right)^p \tag{4.12}$$

and $\log(\zeta_i(\mathbf{r}) - \zeta_i(z))$ as

$$C + \frac{1}{2}\log(\mathbf{r}-z) + \log\left(1 + \sum_{(q,k)\in\mathcal{T}'} \frac{c_{(q,k)}}{c_0}(\mathbf{r}-z)^{q-1/2}\log^k(\mathbf{r}-z) + \mathbf{o}\left(\frac{\max\mathcal{T}'}{\sqrt{\mathbf{r}-z}}\right)\right). \tag{4.13}$$

Once again, if $\max\mathcal{T}' = (\mathbf{r}-z)^{\hat{q}}\log^{\hat{k}}(\mathbf{r}-z)$ then the next possible terms up to order $(\mathbf{r}-z)^{\hat{q}}$ in the expansion may only be

$$(\mathbf{r}-z)^{\hat{q}}\log^{\hat{k}-1}(\mathbf{r}-z), (\mathbf{r}-z)^{\hat{q}}\log^{\hat{k}-2}(\mathbf{r}-z), \ldots, (\mathbf{r}-z)^{\hat{q}}.$$

We determine step by step the corresponding coefficients of these terms by plugging the expansions of $\zeta_i(z)$, (4.12) and (4.13) into Equations (3.1) and (3.2) and comparing error terms. The next term has the form $(\mathbf{r}-z)^{\check{q}}\log^{\check{k}}(\mathbf{r}-z)$, where $\check{q} \leq 2$ is now a sum of elements from the finite set $\{1/2, q/2, q/2 - 1/2 \mid (q,\cdot) \in \mathcal{T}_1 \cup \mathcal{T}_2\}$ such that $\check{q} > \hat{q}$ (recall the definitions of $\mathcal{T}_i$ from (2.2)). Due to (4.12) and (4.13) there is obviously a maximal $\check{k} \in \mathbb{N}_0$ such that $(\mathbf{r}-z)^{\check{q}}\log^{\check{k}}(\mathbf{r}-z)$ may be a non-vanishing next term in the expansion of $\zeta_i(z)$. Iterating the last steps yields the claim of the lemma, since there are only finitely many possible values for q such that the term $(\mathbf{r}-z)^q\log^k(\mathbf{r}-z)$ may appear in the expansion of $\zeta_i(z)$. ∎

Substituting the obtained expansion of $\zeta_1(z)$ into Equation (2.6) yields the proposed claim of Theorem 4.1.

Remark. *The result could also be obtained analogously to Flajolet and Sedgewick* [7, Section VI.7.] *by singularity analysis, but one still has to prove positivity and finiteness of* $\Phi''(\bar{\theta})$.

5. THE REMAINING CASES

In this section we look at all remaining cases not covered by Section 3 and 4. Afterwards we will extend our results to free products $\Gamma_1 * \ldots * \Gamma_m$ with $m > 2$.

5.1. Case $G_1(r_1) < \infty$ and $G_1'(r_1) = \infty$

Theorem 5.1. *Consider a free product of the form* $\Gamma_1 * \Gamma_2$, *where* $G_1(\mathbf{r}_1) < \infty$, $G_1'(\mathbf{r}_1) = \infty$ *and* $G_2'(\mathbf{r}_2) < \infty$. *Then:*

$$\mu^{(n\delta)}(e) \sim \begin{cases} C_1 \cdot \mathbf{r}^{-n\delta} \cdot n^{-3/2}, & \text{if } \bar{\theta} = \theta_1/\alpha_1 \text{ or } \Psi(\bar{\theta}) \leq 0, \\ C_2 \cdot \mathbf{r}^{-n\delta} \cdot n^{-\lambda_2} \cdot \log^{\kappa_2}(n), & \text{if } \bar{\theta} = \theta_2/\alpha_2 < \theta_1/\alpha_1 \text{ and } \Psi(\bar{\theta}) > 0. \end{cases}$$

Proof. For the first part of the proof assume that $\bar{\theta} = \theta_1/\alpha_1$. With

$$U_1(z) := \sum_{g\in\Gamma_1} \mu_1(g)\, z\, F_1(g^{-1}|z)$$

we have the well-known equation $G_1(z) = 1/(1 - U_1(z))$. Therefore, $G_1'(\mathbf{r}_1) = \infty$ implies $U_1'(\mathbf{r}_1) = \infty$, and we get due to [22, Equation (9.14)]

$$\Psi_1(\alpha_1\bar{\theta}) = \Psi_1(\theta_1) = \lim_{z\to\mathbf{r}_1} \Psi_1(zG(z)) = \lim_{z\to\mathbf{r}_1} \frac{1}{zU_1'(z) + 1 - U_1(z)} = 0. \tag{5.1}$$

Thus,

$$\Psi(\bar{\theta}) = \Psi_1(\alpha_1\bar{\theta}) + \Psi_2(\alpha_2\bar{\theta}) - 1 = \Psi_1(\theta_1) + \Psi_2(\alpha_2\bar{\theta}) - 1 = \Psi_2(\alpha_2\bar{\theta}) - 1.$$

Recall that $\Psi(t)$ is strictly decreasing and $\Psi_2(0) = 1$. Therefore, $\Psi(\bar{\theta}) < 0$, and consequently we obtain the asymptotic behaviour $\mu^{(n\delta)}(e) \sim C_1\mathbf{r}^{-n\delta}n^{-3/2}$; see [22, Theorem 17.3].

For the case $\bar{\theta} = \theta_2/\alpha_2 < \theta_1/\alpha_1$ and $\Psi(\bar{\theta}) = 0$, we refer to Section 4.

In the case $\bar{\theta} = \theta_2/\alpha_2 < \theta_1/\alpha_1$ and $\Psi(\bar{\theta}) > 0$ the Green function $G_1(z)$ is analytic at $z = \zeta_1(\mathbf{r}) < \mathbf{r}_1$ and thus we may apply the technique from Section 3 to obtain the proposed asymptotic behaviour. ∎

At this point, let us remark that the formula for $\Psi(t)$ used in Equation (5.1) always implies $\Psi_i(\theta_i) = 0$ whenever $G_i'(\mathbf{r}_i) = \infty$. Moreover:

Corollary 5.2. *If $G_1'(\mathbf{r}_1) = G_2'(\mathbf{r}_2) = \infty$, then $\mu^{(n\delta)}(e) \sim C \cdot \mathbf{r}^{-n\delta} \cdot n^{-3/2}$.*

Proof. Since $U_1'(\mathbf{r}_1) = U_2'(\mathbf{r}_2) = \infty$, Equation (5.1) implies that at least one of $\Psi_1(\alpha_1\bar{\theta})$ and $\Psi_2(\alpha_2\bar{\theta})$ equals zero, yielding $\Psi(\bar{\theta}) < 0$. ∎

5.2. Case $G_1(r_1) = \infty$

For finite groups Γ_1 and Γ_2, Woess [21] proved that the n-step return probabilities behave asymptotically like $C \cdot \mathbf{r}^{-n\delta} \cdot n^{-3/2}$. Moreover, we get the following asymptotic behaviours:

Theorem 5.3. *Consider a free product of the form $\Gamma_1 * \Gamma_2$, where $G_1(\mathbf{r}_1) = \infty$. Then:*

$$\mu^{(n\delta)}(e) \sim \begin{cases} C_1 \cdot \mathbf{r}^{-n\delta} \cdot n^{-3/2}, & \text{if } \Psi(\bar{\theta}) \leq 0, \\ C_2 \cdot \mathbf{r}^{-n\delta} \cdot n^{-\lambda_2} \cdot \log^{\kappa_2}(n), & \text{if } \Psi(\bar{\theta}) > 0. \end{cases}$$

Proof. If $G_2'(\mathbf{r}_2) = \infty$, we have $\Psi(\bar{\theta}) < 0$; see proof of Corollary 5.2.

If $G_2(\mathbf{r}_2) < \infty$ and $G_2'(\mathbf{r}_2) = \infty$ then $\bar{\theta} = \theta_2/\alpha_2$, and $U_2'(\mathbf{r}_2) = \infty$. This implies once again $\Psi(\alpha_2\bar{\theta}) = 0$, and thus $\Psi(\bar{\theta}) < 0$.

If $G_2'(\mathbf{r}_2) < \infty$ then $\bar{\theta} = \theta_2/\alpha_2$ and $\zeta_1(\mathbf{r}) < \mathbf{r}_1$. Therefore, we can follow the argumentation of Section 3 and 4 analogously to prove the proposed claim. ∎

5.3. Free Products with more than two Factors

Let $m \in \mathbb{N}$ with $m \geq 3$. Suppose we are given finitely generated groups $\Gamma_1, \ldots, \Gamma_m$. We consider now a free product of the form $\Gamma := \Gamma_1 * \ldots * \Gamma_m$, on which a random walk is governed by the measure μ defined as $\mu := \sum_{j=1}^{m} \alpha_j\bar{\mu}_j$; see Section 2. We get the following result:

Theorem 5.4. *Let $m \geq 3$. Consider the free product $\Gamma := \Gamma_1 * \ldots * \Gamma_m$ equipped with a random walk governed by $\mu := \sum_{j=1}^{m} \alpha_j \bar{\mu}_j$. Assume that the corresponding Green functions $G_i(z)$ on the free factors Γ_i have an expansion as in (2.2) whenever $G_i'(\mathbf{r}) < \infty$. Denote by $\mathbf{r}$ the radius of convergence of the Green function associated with the random walk on Γ. Then the asymptotic behaviour of the corresponding n-step transition probabilities must obey one of the following laws: $C\,\mathbf{r}^{-n\delta}\, n^{-\lambda_i} \log^{\kappa_i}(n)$, where λ_i and κ_i are inherited from one of the μ_i's, or $C\,\mathbf{r}^{-n\delta}\, n^{-3/2}$ with some constant $C = C_\mu$ depending on μ.*

Proof. In order to prove the theorem, we just remark that – by induction on the number of free factors – the Green function (with radius of convergence $\mathbf{r}^*$) of the random walk on $\Gamma^* := \Gamma_1 * \ldots * \Gamma_{m-1}$ governed by $\mu^* := \sum_{j=1}^{m-1} \frac{\alpha_j}{\alpha_1+\ldots+\alpha_{m-1}} \bar{\mu}_j$ has an expansion either of the form

$$G^*(z) = \sum_{k=0}^{D} g_k(\mathbf{r}^* - z)^k + \sum_{(q,k)\in\mathcal{T}} g_{(q,k)}(\mathbf{r}^* - z)^q \log^k(\mathbf{r}^* - z) + \mathcal{O}((\mathbf{r}^* - z)^{D+2}), \quad \text{(I)}$$

where $\mathcal{T}$ is a finite subset of $\{(q,k) \in \mathbb{R} \times \mathbb{N}_0 \mid D < q \leq D+2\}$ and $g_k, g_{(q,k)} \in \mathbb{R}$, or of the form

$$G^*(z) = g_0 + g_1\sqrt{\mathbf{r}^* - z} + \sum_{(q,k)\in\mathcal{T}} g_{(q,k)}(\mathbf{r}^* - z)^q \log^k(\mathbf{r}^* - z) + \mathcal{O}((\mathbf{r}^* - z)^2), \quad \text{(II)}$$

where $\mathcal{T}$ is a finite subset of $\{(q,k) \in \mathbb{R} \times \mathbb{N}_0 \mid 1/2 < q \leq 2\}$ and $g_0, g_1, g_{(q,k)} \in \mathbb{R}$ with $g_1 \neq 0$. Thus, we may apply the results from Section 3 to the free product $\Gamma^* * \Gamma_m$ equipped with $\mu = (\alpha_1 + \ldots + \alpha_{m-1})\mu^* + \alpha_m \bar{\mu}_m$ and obtain the proposed result. ∎

6. EXAMPLES

6.1. Free Products of Lattices

Let $d_1, \ldots, d_m \in \mathbb{N}$. In this subsection we consider free products of the form $\Gamma := \mathbb{Z}^{d_1} * \ldots * \mathbb{Z}^{d_m}$, equipped with a nearest neighbour random walk, that is, we always assume $\mathrm{supp}(\mu_i) = \{\pm e_j^{(i)} \mid 1 \leq j \leq d_i\}$, where $e_j^{(i)}$ is the j-th unit vector in $\mathbb{Z}^{d_i}$. In the following subsection we show that the Green functions of nearest neighbour random walks on $\mathbb{Z}^d$ have an expansion as requested by (2.2). Afterwards we can give a complete classification of the asymptotic behaviour.

6.1.1. Expansion of the Green Function on $\mathbb{Z}^d$. Let $d \in \mathbb{N}$. Suppose we are given a probability measure π with $\mathrm{supp}(\pi) = \{\pm e_1, \ldots, \pm e_d\}$, the set of natural generators of $\mathbb{Z}^d$. Then π defines a random walk on $\mathbb{Z}^d$, and we denote by $\pi^{(n)}$ its n-fold convolution power. We write for $1 \leq i \leq d$

$$\beta_i := \pi(e_i) + \pi(-e_i) \quad \text{and} \quad p_i := \frac{\pi(e_i)}{\pi(e_i) + \pi(-e_i)}.$$

Denote by $\mathbf{0}$ the zero vector in $\mathbb{Z}^d$. Once again $G_d(z) := \sum_{n\geq 0} \pi^{(n)}(\mathbf{0}) z^n$ denotes the associated Green function, which has radius of convergence $\mathbf{r}_d$. The crucial point for our later discussion is the following:

Proposition 6.1. *The Green function of the random walk on $\mathbb{Z}^d$ has an expansion of the form*

$$G_d(z) = \begin{cases} f(z) + g(z)(\mathbf{r}_d - z)^{(d-2)/2}, & \text{if } d \text{ is odd,} \\ f(z) + g(z)(\mathbf{r}_d - z)^{(d-2)/2}\log(\mathbf{r}_d - z), & \text{if } d \text{ is even,} \end{cases}$$

where the functions $f(z), g(z)$ are analytic in a neighbourhood of $z = \mathbf{r}_d$ and $g(\mathbf{r}_d) \neq 0$.

Remarks. For the case of simple random walks on $\mathbb{Z}^d$, i.e. $\pi(\pm e_i) = 1/(2d)$, a proof of this proposition can be found in [22, Proposition 17.16]. In our case, we generalize the statement to arbitrary nearest neighbour random walks on $\mathbb{Z}^d$, but we will only give a sketch of the proof and refer once again to [22]. From the expansion follows with the help of Darboux's method that $\pi^{(2n)}(\mathbf{0}) \sim C\,\mathbf{r}_d^{-2n}\,n^{-d/2}$; this asymptotic behaviour follows also from Cartwright and Soardi [4].

Proof. First, note that the spectral radius of the random walk on $\mathbb{Z}^d$ is given by

$$\varrho = \sum_{i=1}^{d} \beta_i \sqrt{4p_i(1-p_i)} = \frac{1}{\mathbf{r}_d};$$

compare with [22, Theorem 8.23]. For $i \in \{1, \dots, d\}$, we define random walks on $\mathbb{Z}$ governed by probability measures π_i with $\pi_i(1) := p_i$ and $\pi_i(-1) := 1-p_i$. For $z \in \mathbb{C}$, the *exponential generating function* on $\mathbb{Z}^d$ is given by

$$E(z) := \sum_{n=0}^{\infty} \pi^{(n)}(\mathbf{0}) \frac{z^n}{n!}$$

and on the i-th coordinate axis it is given by

$$E_i(z) := \sum_{n\geq 0} \pi_i^{(n)}(0) \frac{z^n}{n!} = \int_{-1}^{1} e^{\sqrt{4p_i(1-p_i)}tz} \frac{1}{\pi\sqrt{1-t^2}} dt.$$

In the last equation we applied the following relation, which is easy to check:

$$\pi_i^{(n)}(0) = \int_{-1}^{1} \sqrt{4p_i(1-p_i)}^{\,n}\, t^n \frac{1}{\pi\sqrt{1-t^2}} dt.$$

Furthermore, we get $E(z) = \prod_{i=1}^{d} E_i(\beta_i z) = \int_{-\varrho}^{\varrho} e^{tz} (\hat{f}_1 * \dots * \hat{f}_d)(t) dt$, where

$$\hat{f}_i(t) := \frac{1}{\beta_i\sqrt{4p_i(1-p_i)}} f_0\left(\frac{t}{\beta_i\sqrt{4p_i(1-p_i)}}\right) \quad \text{and} \quad f_0(t) := \begin{cases} \frac{1}{\pi\sqrt{1-t^2}}, & \text{if } t \in (-1,1), \\ 0, & \text{otherwise.} \end{cases}$$

This allows us to rewrite the Green function in the following way:

$$G_d(z) = \int_{-\varrho}^{\varrho} \frac{1}{1-zt} (\hat{f}_1 * \dots * \hat{f}_d)(t)\, dt. \tag{6.1}$$

Moreover, there is a function $g_d(t)$, which is analytic in a neighbourhood of $t = \varrho$ and satisfies $g_d(\varrho) \neq 0$ such that

$$(\hat{f}_1 * \ldots * \hat{f}_d)(t) = (\varrho - t)^{(d-2)/2} g_d(t). \tag{6.2}$$

To prove this, we define $\bar{f}_i(t) := \hat{f}_i(\beta_i\sqrt{4p_i(1-p_i)} - t)$ and show inductively that we can write

$$(\bar{f}_1 * \ldots * \bar{f}_d)(t) = t^{(d-2)/2} \bar{g}_d(t),$$

where the function $\bar{g}_d(t)$ is analytic in a neighbourhood of $t = 0$ and $\bar{g}_d(0) \neq 0$. Analogously to the proof of [22, Proposition 17.16], we may conclude together with (6.1) and (6.2) that $G_d(z)$ has the proposed expansion. ∎

6.1.2. Classification of the Asymptotic Behaviour. Observe that a nearest neighbour random walk on $\mathbb{Z}^d$ has period 2 since it can return to the origin only in an even number of steps. Therefore, the period of a nearest neighbour random walk on $\mathbb{Z}^{d_1} * \mathbb{Z}^{d_2}$ is $\delta = 2$. Now we can give a complete classification of the asymptotic behaviour of n-step return probabilities of nearest neighbour random walks on $\mathbb{Z}^{d_1} * \mathbb{Z}^{d_2}$:

Theorem 6.2. *Consider irreducible nearest neighbour random walks on the lattices $\mathbb{Z}^{d_1}$ and $\mathbb{Z}^{d_2}$ with $d_1 \leq d_2$. Then the n-step return probabilities of the associated random walk on $\mathbb{Z}^{d_1} * \mathbb{Z}^{d_2}$ obey one the following laws:*

$$\mu^{(2n)}(e) \sim \begin{cases} C_1 \cdot \mathbf{r}^{-2n} \cdot n^{-d_1/2}, & \text{if } d_1 \geq 5 \textit{ and } \Psi(\bar{\theta}) > 0 \textit{ and } \bar{\theta} = \theta_1/\alpha_1, \\ C_2 \cdot \mathbf{r}^{-2n} \cdot n^{-d_2/2}, & \text{if } d_2 \geq 5 \textit{ and } \Psi(\bar{\theta}) > 0 \textit{ and } \bar{\theta} = \theta_2/\alpha_2 < \theta_1/\alpha_1, \\ C_3 \cdot \mathbf{r}^{-2n} \cdot n^{-3/2}, & \textit{otherwise}. \end{cases}$$

Consider now the multi-factor free product $\mathbb{Z}^{d_1} * \ldots * \mathbb{Z}^{d_m}$. Let μ_i be the simple random walk on $\mathbb{Z}^{d_i}$ for each $i \in \{1, \ldots, m\}$, that is, $\mu_i(\pm e_j^{(i)}) = 1/(2d_i)$, where $e_j^{(i)}$ is the j-th unit vector in $\mathbb{Z}^{d_i}$. Choose $\alpha_1, \ldots, \alpha_m > 0$ with $\sum_{j=1}^m \alpha_j = 1$. Let $G_i(z)$ denote the Green function of the simple random walk on $\mathbb{Z}^{d_i}$, which has radius of convergence $\mathbf{r}_i = 1$, and define $\Psi_i(t)$ analogously as in (2.8). Cartwright [1] computed numerically some of the values of $\Psi_i(G_i(1))$ and showed that $\Psi_i(G_i(1)) \to 1$ if $d_i \to \infty$. Thus, for large d_i we have $\Psi_i(G_i(1)) > 1 - 1/m$. Recall also that $\Psi_i(t)$ is decreasing. Denote by $G(z)$ the Green function of the random walk on $\mathbb{Z}^{d_1} * \ldots * \mathbb{Z}^{d_m}$ and by $\mathbf{r}$ its radius of convergence, and define $\Psi(t)$ analogously as in (2.8). By [22, Equation 9.21],

$$\Psi(\bar{\theta}) = 1 + \sum_{j=1}^{m} (\Psi_i(\alpha_i\bar{\theta}) - 1),$$

where $\bar{\theta} = \min_{1 \leq i \leq m} \theta_i/\alpha_i$. If all exponents $d_i \geq 5$ are large enough, we get $\Psi(\bar{\theta}) > 0$. Furthermore, if α_i is chosen large enough, we get an asymptotic behaviour of the form $C_i\, \mathbf{r}^{-2n}\, n^{-d_i/2}$. Moreover, one can define (symmetric) measures $\mu_1, \ldots, \mu_m$ supported on the natural generators in such a way that we obtain a $C_0\, \mathbf{r}^{-2n}\, n^{-3/2}$-law: one chooses μ_1 and μ_2

such that $\Psi_1(\theta_1), \Psi_2(\theta_2) < 1/2$, and α_1 and α_2 are chosen such that $\bar{\theta} = \theta_1/\alpha_1 = \theta_2/\alpha_2$, yielding

$$\Psi(\bar{\theta}) = 1 + \underbrace{(\Psi_1(\theta_1) - 1)}_{<-1/2} + \underbrace{(\Psi_2(\theta_2) - 1)}_{<-1/2} + \underbrace{\sum_{k=3}^{m}(\Psi_k(\alpha_k\bar{\theta}) - 1)}_{\leq 0} < 0;$$

see also comments at the end of Section 2. That is, we can have $m+1$ different asymptotic behaviours. This finally proves Theorem 1.1.

For instance, consider $\Gamma = \mathbb{Z}^5 * \mathbb{Z}^6 * \mathbb{Z}^7$ equipped with simple random walks μ_1, μ_2 and μ_3 on each free factor. For $i \in \{1,2,3\}$, we define $\Psi_i(t)$ analogously to (2.8). Cartwright [1] computed the values $\Psi_1(G_1(1)) = 0.691$, $\Psi_2(G_2(1)) = 0.824$ and $\Psi_3(G_3(1)) = 0.876$. Thus, the random walk on $\mathbb{Z}^5 * \mathbb{Z}^6$ governed by $\mu_{12} := \alpha_1^*\bar{\mu}_1 + \alpha_2^*\bar{\mu}_2$, where $\alpha_1^* = \alpha_1/(\alpha_1+\alpha_2)$ and $\alpha_2^* = \alpha_2/(\alpha_1+\alpha_2)$, satisfies $\Psi(M) \geq 0.515$ with $M := \min\{\theta_1/\alpha_1^*, \theta_2/\alpha_2^*\}$. That is, $M = \mathbf{r}_{1,2}G_{1,2}(\mathbf{r}_{1,2})$, where $G_{1,2}(z)$ is the Green function of the random walk on $\mathbb{Z}^5 * \mathbb{Z}^6$ with radius of convergence $\mathbf{r}_{1,2}$. Since all Ψ_i-functions are strictly decreasing, we obtain for the random walk on $\Gamma = \Gamma_1 * \Gamma_2$ with $\Gamma_1 = \mathbb{Z}^5 * \mathbb{Z}^6$ and $\Gamma_2 = \mathbb{Z}^7$:

$$\Psi(\bar{\theta}) = \Psi_1((\alpha_1+\alpha_2)\bar{\theta}) + \Psi_2(\alpha_3\bar{\theta}) - 1 \geq 0.515 + 0.876 - 1 > 0.$$

For the simple random walk on Γ, we have then the asymptotic non-exponential type $n^{-7/2}$, if $\alpha_1 + \alpha_2 < M/(M + G_3(1))$. Otherwise, we have the asymptotic behaviour $n^{-5/2}$, if $M = \theta_1/\alpha_1^*$, or n^{-3}, if $M = \theta_2/\alpha_2^* \neq \theta_1/\alpha_1^*$.

6.2. $(\mathbb{Z}/m\mathbb{Z}) * \mathbb{Z}^d$

Consider the groups $\Gamma_1 = \mathbb{Z}/m\mathbb{Z}$ and $\Gamma_2 = \mathbb{Z}^d$ for any $m, d \in \mathbb{N}$ with $m \geq 2$. Suppose we are given a probability measure μ_1 on Γ_1 and a probability measure μ_2 on $\mathbb{Z}^d$, which is supported on the natural generators. Then $G_1(1) = \infty$, and thus we get the following classification:

$$\mu^{(n\delta)}(e) \sim \begin{cases} C_1 \cdot \mathbf{r}^{-n\delta} \cdot n^{-d/2}, & \text{if } \Psi(\bar{\theta}) > 0, \\ C_2 \cdot \mathbf{r}^{-n\delta} \cdot n^{-3/2}, & \text{otherwise.} \end{cases}$$

Let us remark that $\Psi(\bar{\theta}) < 0$ if $d \leq 4$: this follows from the fact $G_2'(\mathbf{r}_2) = \infty$ (see Proposition 6.1) and Corollary 5.2.

6.3. $\Pi_q * \mathbb{Z}^d$

Consider the groups $\Gamma_1 = \Pi_q := *_{i=1}^{q}(\mathbb{Z}/2\mathbb{Z})$ and $\Gamma_2 = \mathbb{Z}^d$ for any $q, d \in \mathbb{N}$ with $q \geq 2$. Observe that the Cayley graph of Γ_1 is the homogeneous tree of degree q. Suppose we are given probability measures μ_1 on Γ_1 and μ_2 on $\mathbb{Z}^d$, which are both supported on the natural generators. If $q = 2$ then $G_1(1) = \infty$, and thus we get the same classification as in the case $(\mathbb{Z}/m\mathbb{Z}) * \mathbb{Z}^d$. If $q \geq 3$, then it is well-known that $G_1(z)$ can be written as

$$G_1(z) = A(z) + \sqrt{\mathbf{r}_1 - z}\,B(z),$$

where $A(z), B(z)$ are analytic in a neighbourhood of $z = \mathbf{r}_1$ and $B(\mathbf{r}_1) \neq 0$; see e.g. Woess [23, Equation (4.5)]. Therefore, we get the following classification for the associated random walk on the free product $\Gamma_1 * \Gamma_2$:

$$\mu^{(2n)}(e) \sim \begin{cases} C_1 \cdot \mathbf{r}^{-2n} \cdot n^{-d/2}, & \text{if } \bar{\theta} = \theta_2/\alpha_2 < \theta_1/\alpha_1 \text{ and } \Psi(\bar{\theta}) > 0, \\ C_2 \cdot \mathbf{r}^{-2n} \cdot n^{-3/2}, & \text{otherwise.} \end{cases}$$

Analogously to the previous example, observe that $d \leq 4$ implies $\Psi(\bar{\theta}) < 0$.

7. CLASSIFICATION OF PHASE TRANSITIONS

Let us return to the case $m = 2$, that is, $\Gamma = \Gamma_1 * \Gamma_2$. We now *fix* the measures μ_1 and μ_2, and investigate the variation of $\Psi(\bar{\theta})$ as a function of the parameter α_1.

Lemma 7.1. *Assume $\bar{\theta} < \infty$. Then the function $\Upsilon : (0, 1) \mapsto \mathbb{R}$ defined by*

$$\Upsilon(\alpha_1) := \Psi_1(\alpha_1\bar{\theta}) + \Psi_2((1-\alpha_1)\bar{\theta}) - 1$$

is continuous, strictly decreasing in the interval $(0, \frac{\theta_1}{\theta_1+\theta_2}]$ and strictly increasing in the interval $[\frac{\theta_1}{\theta_1+\theta_2}, 1)$. (We set $\frac{c}{c+\infty} := 0$ and $\frac{\infty}{\infty+c} := 1$ for $c \in (0, \infty)$.)

Proof. We leave the proof of continuity of Υ as an easy exercise to the reader, since Ψ_i is analytic in an open neighbourhood of the interval $[0, \theta_i)$.

Note that $\Upsilon(\alpha_1)$ equals $\Psi(\bar{\theta})$ in dependence of α_1. We divide the proof into two parts, according to finiteness of θ_1 and θ_2.

Case $\theta_1, \theta_2 < \infty$. If $0 < \alpha_1 < \frac{\theta_1}{\theta_1+\theta_2}$ then $\bar{\theta} = \theta_2/\alpha_2$. Consequently, we have

$$\Upsilon(\alpha_1) = \Psi_1\left(\frac{\alpha_1}{1-\alpha_1}\theta_2\right) + \Psi_2(\theta_2) - 1.$$

Since the function $\frac{\alpha_1}{1-\alpha_1}$ is strictly increasing, it follows that $\Psi_1(\frac{\alpha_1}{1-\alpha_1}\theta_2)$ is strictly decreasing, implying $\Upsilon(\alpha_1)$ strictly decreasing.

If $\alpha_1 = \frac{\theta_1}{\theta_1+\theta_2}$ we obtain $\bar{\theta} = \theta_1/\alpha_1 = \theta_2/\alpha_2$, that is, $\Upsilon(\alpha_1) = \Psi_1(\theta_1) + \Psi_2(\theta_2) - 1$.

If $\frac{\theta_1}{\theta_1+\theta_2} < \alpha_1 < 1$ we have $\Psi(\bar{\theta}) = \Psi_1(\theta_1) + \Psi_2(\frac{1-\alpha_1}{\alpha_1}\theta_1) - 1$. Since $\frac{1-\alpha_1}{\alpha_1}$ is strictly decreasing, $\Upsilon(\alpha_1)$ is a strictly increasing function in the abovementioned interval.

Case $\theta_1 = \infty$. Then $\bar{\theta} = \frac{\theta_2}{1-\alpha_1}$. The same reasoning as before shows that $\Upsilon(\alpha_1)$ is strictly decreasing in the interval $(0, 1)$.

Case $\theta_2 = \infty$. Then $\bar{\theta} = \frac{\theta_1}{\alpha_1}$. Analogously, $\Upsilon(\alpha_1)$ is strictly increasing in the interval $(0, 1)$. ∎

Let us remark that $\bar{\theta} = \infty$ implies $\Psi(\bar{\theta}) < 0$ (see [22, Theorem 9.22]); otherwise we would have a contradiction to ρ-transience.

Now we can give a complete picture of the *phase transition* of the asymptotic behaviour of the return probabilities depending on the parameter α_1, and we present specific examples. In the following we discuss the different possible behaviours of the function $\Upsilon(\alpha_1) = \Psi(\bar{\theta})$. In Figure 1, the dashed line will represent *approximately* the qualitative behaviour of $\Upsilon(\alpha_1)$; we denote its zeros (if they exist) by α_{low} and α_{high} (with $\alpha_{\text{low}} \leq \alpha_{\text{high}}$). Moreover, we write

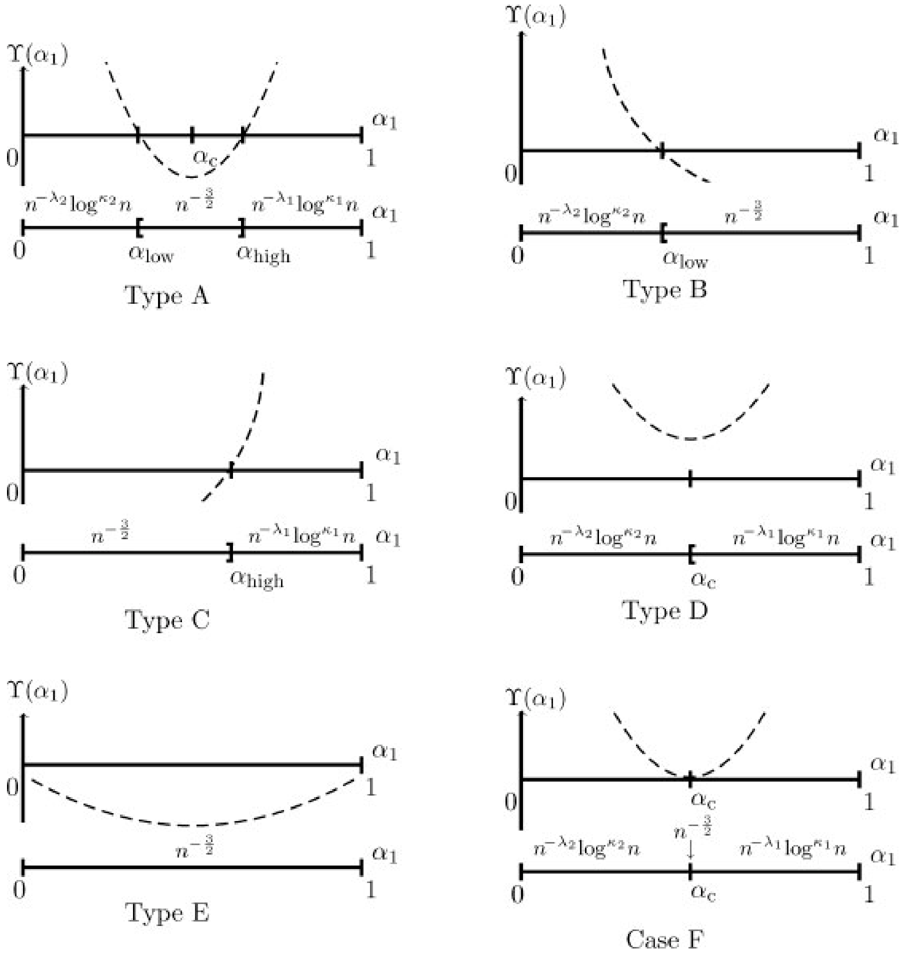


Fig. 1. The different behaviours of $\Upsilon : \alpha_1 \mapsto \Psi(\bar{\theta})$.

$\alpha_c := \theta_1/(\theta_1+\theta_2)$. We decompose the interval (0, 1) into subintervals such that every choice of α_1 in a fixed subinterval leads to the same non-exponential type. With the help of Figure 1 we discuss case by case the different behaviours of $\Upsilon(\alpha_1)$, and for each case we give an example of a nearest neighbour random walk on $\mathbb{Z}^{d_1} * \mathbb{Z}^{d_2}$. Recall that $\Psi(0) = \Psi_i(0) = 1$.

Case A: Consider Figure 1, Case A. We give an example such that this case holds. We set $\Gamma = \mathbb{Z}^{d_1} * \mathbb{Z}^{d_2}$ with $d_1, d_2 \geq 5$, and we choose μ_1 and μ_2 such that $\Psi_1(\theta_1) < 1/2$ and $\Psi_2(\theta_2) < 1/2$. Recall that it is possible to find such measures (see end of Section 2 and [22, Lemma 17.9]). We remark that $\Psi_i(\theta_i) > 0$: indeed, $\Psi_i(\theta_i) = 0$ would imply

$$\Phi_i'(\theta_i) = \frac{\Phi_i(\theta_i)}{\theta_i} = \frac{G_i(\mathbf{r}_i)}{\mathbf{r}_i G_i(\mathbf{r}_i)} = \frac{1}{\mathbf{r}_i}.$$

Differentiating (2.7) would yield $G_i'(\mathbf{r}_i) = \infty$, a contradiction to Proposition 6.1, according to which, $G_i'(\mathbf{r}_i)$ must be finite due to $d_i \geq 5$. Therefore:

– If α_1 is small then $\bar{\theta} = \theta_2/(1-\alpha_1)$ and

$$\Psi(\bar{\theta}) = \underbrace{\Psi_1\Bigg(\underbrace{\alpha_1 \frac{\theta_2}{1-\alpha_1}}_{\xrightarrow{\alpha_1 \to 0} 0}\Bigg)}_{\xrightarrow{\alpha_1 \to 0} 1} + \underbrace{\Psi_2(\theta_2)}_{>0} - 1, \tag{7.1}$$

that is, $\Psi(\bar{\theta}) > 0$ if α_1 is sufficiently small. This yields an $n^{-d_2/2}$-law for small values of α_1.

– If α_1 is close to 1 then $\bar{\theta} = \theta_1/\alpha_1$ and we get analogously an $n^{-d_1/2}$-law.

– For $\alpha_1 = \alpha_c$, we get $\Psi(\bar{\theta}) = \Psi_1(\theta_1) + \Psi_2(\theta_2) - 1 < 0$, that is, we have an $n^{-3/2}$-law in this case.

Case B: We set $\Gamma = \mathbb{Z}^2 * \mathbb{Z}^7$. By Lemma 7.1, $\Upsilon(\alpha_1)$ is strictly decreasing and $\bar{\theta} = \theta_2/\alpha_2$.

– If α_1 is small then the same reasoning as in (7.1) holds and $\Psi(\bar{\theta}) > 0$, that is, we have an $n^{-d_2/2}$-law for small α_1.

– If α_1 is close to 1 then

$$\Psi(\bar{\theta}) = \underbrace{\Psi_1\Bigg(\underbrace{\alpha_1 \frac{\theta_2}{1-\alpha_1}}_{\xrightarrow{\alpha_1 \to 1} \infty}\Bigg)}_{\xrightarrow{\alpha_1 \to 1} 0} + \underbrace{\Psi_2(\theta_2)}_{<1} - 1 < 0,$$

since $\lim_{t\to\infty} \Psi_1(t) = 0$, which follows analogously to (5.1). That is, we have an $n^{-3/2}$-law for large α_1.

Case C: By setting $\Gamma = \mathbb{Z}^7 * \mathbb{Z}^2$, we have the symmetric situation as in Case B, which gives an example for this case by exchanging the roles of $\mathbb{Z}^2$ and $\mathbb{Z}^7$.

Case D: We set $\Gamma = \mathbb{Z}^5 * \mathbb{Z}^6$ and consider simple random walks on the factors $\mathbb{Z}^5$ and $\mathbb{Z}^6$. By Cartwright [1], we have $\Psi_1(\theta_1) = 0.691$ and $\Psi_2(\theta_2) = 0.824$. Since $\Psi_1(z)$ and $\Psi_2(z)$ are strictly decreasing, we have $\Upsilon(\alpha_1) \geq \Psi_1(\theta_1) + \Psi_2(\theta_2) - 1 > 0$ for all $\alpha_1 \in (0,1)$. Thus, we obtain an $n^{-5/2}$-law, if $\alpha_1 \geq \alpha_c$, and an n^{-3}-law, if $\alpha_1 < \alpha_c$.

Case E: We set $\Gamma = \mathbb{Z}^3 * \mathbb{Z}^4$. By Equation (5.1), follows that $\Psi_1(\alpha_1\bar{\theta}) = 0$ or $\Psi_2(\alpha_2\bar{\theta}) = 0$, that is, we have $\Upsilon(\alpha_1) < 0$ for all $\alpha_1 \in (0,1)$. This yields an $n^{-3/2}$-law for all $\alpha_1 \in (0,1)$.

We now give an example (see Case F of Figure 1) where the $n^{-3/2}$-interval of case A collapses to a singleton. For this purpose, we have to prove the following:

Lemma 7.2. *Consider $\Gamma = \mathbb{Z}^5 * \mathbb{Z}^6$. Then there are probability measures μ_1 and μ_2 supported on the natural generators of $\mathbb{Z}^5$ and $\mathbb{Z}^6$ respectively such that $\Psi_1(\theta_1) = \Psi_2(\theta_2) = \frac{1}{2}$.*

Proof. Let $i \in \{1,2\}$. We have $d_1 = 5, d_2 = 6$ and choose any $\delta \in (0,1)$. We define

$$\nu_\delta^{(i)}(x) := \begin{cases} (1-\delta)/2, & \text{if } x = (\pm 1, 0, \dots, 0) \in \mathbb{Z}^{d_i}, \\ \dfrac{\delta}{2d_i - 2}, & \text{if } x = (0, \dots, 0, \pm 1, 0, \dots, 0) \in \mathbb{Z}^{d_i} \setminus \{(\pm 1, 0, \dots, 0)\}. \end{cases}$$

The Green function associated with the random walk on $\mathbb{Z}^{d_i}$ governed by the symmetric measure $\nu_\delta^{(i)}$ has radius of convergence $\mathbf{r}_i = 1$; see [22, Cor. 8.15]. If $\delta = 1 - 1/d_i$ then $\Psi_1(\theta_1) = 0.691 > 1/2$ and $\Psi_2(\theta_2) = 0.824 > 1/2$; see Cartwright [1]. On the other hand side, if δ is small enough then $\Psi_1(\theta_1) < 1/2$ and $\Psi_2(\theta_2) < 1/2$; see proof of [22, Lemma 17.9]. It remains to show that $\Psi_i(\theta_i)$ varies continuously in dependence of δ, which implies that there is some $\delta_0^{(i)}$ such that $\Psi_i(\theta_i) = 1/2$. We now write $G_i(z) = G_i(\delta|z), U_i(z) = U_i(\delta|z)$ and $\Psi_i(t) = \Psi_i(\delta|t)$. Recall that

$$\Psi_i(\delta|\theta_i) = \frac{1}{U_i'(\delta|1) + 1 - U_i(\delta|1)}.$$

Since $U_i(\delta|1)$ can be rewritten as a power series in the variable δ, the function $\delta \mapsto \Psi_i(\delta|\theta_i)$ is continuous in δ. This finishes the proof. ∎

We can now present an example, where Case F of Figure 1 holds: we set $\Gamma = \mathbb{Z}^5 * \mathbb{Z}^6$ and choose the measures μ_1 and μ_2 such that $\Psi_1(\theta_1) = \Psi_2(\theta_2) = 1/2$. Obviously, we have then $\Upsilon(\alpha_c) = \Psi_1(\theta_1) + \Psi_2(\theta_2) - 1 = 0$. That is, we get the following asymptotic behaviour:

$$\mu^{(2n)}(e) \sim \begin{cases} C_1 \cdot \mathbf{r}^{-2n} \cdot n^{-5/2}, & \text{if } \alpha_1 > \alpha_c, \\ C_2 \cdot \mathbf{r}^{-2n} \cdot n^{-3/2}, & \text{if } \alpha_1 = \alpha_c, \\ C_3 \cdot \mathbf{r}^{-2n} \cdot n^{-3}, & \text{if } \alpha_1 < \alpha_c. \end{cases}$$

As a final remark let us explain that it is not possible that $\Upsilon(\alpha_1)$ is strictly increasing or decreasing with $\Upsilon(\alpha_1) > 0$ for all $\alpha_1 \in (0,1)$. In order to show this assume that $\Upsilon(\alpha_1)$ is strictly increasing. Then, by Lemma 7.1, $\theta_2 = \infty$ must hold, that is, $G_2(\mathbf{r}_2) = \infty$. The same reasoning as in Equation (5.1) leads to $\lim_{z \to \mathbf{r}_2} \Psi_2(zG(z)) = \lim_{t \to \infty} \Psi_2(t) = 0$. Therefore, we obtain for α_1 small enough

$$\Psi(\bar{\theta}) = \underbrace{\Psi_1(\theta_1)}_{<1} + \underbrace{\Psi_2\underbrace{\left((1-\alpha_1)\frac{\theta_1}{\alpha_1}\right)}_{\xrightarrow{\alpha_1 \to 0} \infty}}_{\xrightarrow{\alpha_1 \to 0} 0} - 1 < 0.$$

Analogously, if $\Upsilon(\alpha_1)$ is strictly decreasing, then it must have a zero.

8. HIGHER ASYMPTOTIC ORDERS

The techniques we used for determining the asymptotic behaviour give us not only the leading term $n^{-\lambda}\log^{\kappa} n$, but also the proceeding terms of *higher order*, according to the singular terms in the expansion following the leading one. For instance, consider a nearest neighbour random walk on $\mathbb{Z}^7 * \mathbb{Z}^8$ with $\alpha_1 = \theta_1/(\theta_1+\theta_2)$. Then the associated Green function has the following expansion:

$$\sum_{k=0}^{4} g_k(\mathbf{r}-z)^4 + \hat{g}_1(\mathbf{r}-z)^{5/2} + \check{g}_1(\mathbf{r}-z)^3\log(\mathbf{r}-z) + \hat{g}_2(\mathbf{r}-z)^{7/2} + \check{g}_2(\mathbf{r}-z)^4\log(\mathbf{r}-z) + \mathbf{o}((\mathbf{r}-z)^4),$$

where $\hat{g}_1 \neq 0$. That is,

$$\mu^{(2n)}(e) \sim \mathbf{r}^{-2n}\cdot\left(C_1 n^{-7/2} + C_2 n^{-4} + C_3 n^{-9/2} + C_4 n^{-5} + \mathbf{o}(n^{-5})\right),$$

where $C_1 \neq 0$.

REFERENCES

[1] D. Cartwright, Some examples of random walks on free products of discrete groups, Ann Mat Pura Appl 151 (1988), 1–15.

[2] D. Cartwright, On the asymptotic behaviour of convolution powers of probabilities on discrete groups, Monatsh Math 107 (1989), 287–290.

[3] D. Cartwright and P. Soardi, Random walks on free products, quotients and amalgams, Nagoya Math J 102 (1986), 163–180.

[4] D. Cartwright and P. Soardi, A local limit theorem for random walks on the cartesian product of discrete groups, Boll Unione Mat Ital A VII 1 (1987), 107–115.

[5] I. Chatterji, C. Pittet, and L. Saloff-Coste, Connected Lie groups and property RD, Duke Math J 137 (2007), 511–536.

[6] P. Flajolet and A. Odlyzko, Singularity analysis of generating functions, SIAM J Alg Discr Meth 3 (1990), 216–240.

[7] P. Flajolet and R. Sedgewick, Analytic combinatorics, Cambridge University Press, Cambridge, 2009.

[8] P. Gerl, A local central limit theorem on some groups, In The first pannonian symposium on mathematical statistics, P. Revesz (Editor), Lect. Notes Statistics 8, Springer, 1981, pp. 73–82.

[9] P. Gerl and W. Woess, Local limits and harmonic functions for nonisotropic random walks on free groups, Probab Theory Relat Fields 71 (1986), 341–355.

[10] L. A. Gilch, Rate of escape of random walks on free products, J Aust Math Soc 83 (2007), 31–54.

[11] L. A. Gilch, Asymptotic entropy of random walks on free products, Electron J Probab 16 (2011), 76–105.

[12] S. Lalley, Finite range random walk on free groups and homogeneous trees, Ann Probab 21 (1993), 2087–2130.

[13] S. Lalley, Random walks on infinite free products and infinite algebraic systems of generating functions, preprint, 2002.

[14] R. Lyndon and P. Schupp, Combinatorial group theory, Springer-Verlag, Berlin, 1977.

[15] J. Mairesse and F. Mathéus, Random walks on free products of cyclic groups, J Lond Math Soc 75 (2007), 47–66.

[16] J. McLaughlin, Random walks and convolution operators on free products, Ph.D. Thesis, New York Univ., 1986.

[17] F. Olver, Asymptotics and special functions, Academic Press, San Diego, California, 1974.

[18] S. Sawyer, Isotropic random walks in a tree, Z Wahrscheinlichkeitstheorie Verw Geb 42 (1978), 279–292.

[19] D. Voiculescu, Addition of certain non-commuting random variables, J Funct Anal 66 (1986), 323–346.

[20] W. Woess, A local limit theorem for random walks on certain discrete groups, In Probability measures on groups (Oberwolfach, 1981), H. Heyer (Editor), volume 928 of Lecture Notes in Math., Springer, Berlin, 1982, pp. 467–477.

[21] W. Woess, Nearest neighbour random walks on free products of discrete groups, Boll Unione Mat Ital 5-B (1986), 961–982.

[22] W. Woess, Random walks on infinite graphs and groups, Cambridge University Press, Cambridge, 2000.

[23] W. Woess, Generating function techniques for random walks on graphs, in “Heat Kernels and Analysis on Manifolds, Graphs, and Metric Spaces”, Contemp Math 338 (2003), 391–423.

Publication C

Branching Random Walks on Free Products of Groups

E. Candellero, L. Gilch and S. Müller

Proceedings of the London Mathematical Society,
(3) 104, no. 6, pages 1085–1120, 2012.

Proceedings of the London Mathematical Society Advance Access published January 29, 2012

Proc. London Math. Soc. Page 1 of 36
doi:10.1112/plms/pdr060

Branching random walks on free products of groups

Elisabetta Candellero, Lorenz A. Gilch and Sebastian Müller

Abstract

We study certain phase transitions of branching random walks (BRW) on Cayley graphs of free products. The aim of this paper is to compare the size and structural properties of the trace, that is, the subgraph that consists of all edges and vertices that were visited by some particle, with those of the original Cayley graph. We investigate the phase when the growth parameter λ is small enough such that the process survives, but the trace is not the original graph. A first result is that the box-counting dimension of the boundary of the trace exists, is almost surely constant and equals the Hausdorff dimension which we denote by $\Phi(\lambda)$. The main result states that the function $\Phi(\lambda)$ has only one point of discontinuity which is at $\lambda_c = R$ where R is the radius of convergence of the Green function of the underlying random walk. Furthermore, $\Phi(R)$ is bounded by one half the Hausdorff dimension of the boundary of the original Cayley graph and the behaviour of $\Phi(R) - \Phi(\lambda)$ as $\lambda \uparrow R$ is classified.

In the case of free products of infinite groups the end-boundary can be decomposed into words of finite and words of infinite length. We prove the existence of a phase transition such that if $\lambda \leqslant \tilde{\lambda}_c$, the end boundary of the trace consists only of infinite words and if $\lambda > \tilde{\lambda}_c$, it also contains finite words. In the last case, the Hausdorff dimension of the set of ends (of the trace and the original graph) induced by finite words is strictly smaller than the one of the ends induced by infinite words.

Contents

1. *Introduction*

A branching random walk (BRW) is a *growing cloud* of particles that move on an underlying graph $\mathcal{X}$ in discrete time. The process starts with one particle in the root e of the graph. Then at each discrete time step a particle produces offspring particles according to some offspring distribution with mean $\lambda > 1$, and then each descendent moves one step according to a random walk on $\mathcal{X}$. Particles branch and move independently of the other particles and the history of the process. A first natural question is to ask whether the process eventually *fills up* the whole graph, that is, whether every finite subset will eventually be occupied or free of particles. If the BRW visits the whole graph, it is called recurrent, and transient otherwise. As a consequence

Received 11 May 2011.

2010 *Mathematics Subject Classification* 60J10, 60J80 (primary), 37F35, 20E06 (secondary).

The research of the first author was financially supported by the Austrian Academy of Science (ÖAW) and by the Austrian Science Fund (FWF): W1230-N13, the research of the second author was supported by the German Research Foundation (DFG) grant GI 746/1-1, while the third author was supported by PIEF-GA-2009-235688. Part of the work was done during a visit of S. Müller at Graz University of Technology and a visit of E. Candellero at Geneva University, where both visits were supported by the ESF Grant 'Random Geometry of Large Interacting Systems and Statistical Physics'.

of Kesten's amenability criterion, any BRW is recurrent on the Cayley graph of an amenable group. Furthermore, one observes a phase transition on non-amenable groups; there exists some $\lambda_c > 1$ such that a BRW with $\lambda \leqslant \lambda_c$ is transient, while it is recurrent otherwise. In the transient case, the trace of the BRW, that is, the subgraph that consists of all edges and vertices that were visited by the BRW, is a proper random subgraph of the original Cayley graph. Benjamini and Müller [1] studied first general qualitative statements of the trace of BRW on groups. In particular, they proved exponential volume growth of the trace in general. However, their approach is rather qualitative and gives no quantitative results on the growth rate. In this article, we study BRW on free products of groups and obtain a precise formula for the growth rate and dimensions of the end boundary of the trace. One motivation to study BRW on this class of structures lies mainly in the fact that they are among the simplest non-amenable groups. This makes them to a reference and starting point for more complicated non-amenable structures such as, for instance, groups with infinitely many ends or hyperbolic groups. Besides this, free products of groups are interesting on their own since they play an important role in some fields of algebraic topology and in Basse–Serre theory.

The starting point of the present investigation of BRWs was the work of Hueter and Lalley [13], who studied BRW on homogeneous trees. We remark that in their setting and notation weak survival is equivalent to transience in our language. In the transient regime, the BRW eventually vacates every finite subset and the particle trails converge to the geometric end boundary Ω of the tree. The *limit set* Λ of the BRW is the random subset of the boundary that consists of all ends, where the BRW accumulates. By this we mean that each neighbourhood of an end in Λ is visited infinitely often by the process. Equivalently, we can define Λ as the geometric end boundary of the trace.

Typical ways of measuring the size of boundaries are by use of the box-counting dimension (also known as the Minkowski dimension) or the Hausdorff dimension. In [13], a formula for the Hausdorff dimension of Λ is given for BRW on homogeneous trees. In particular, it is shown there that the limit set has the Hausdorff dimension no larger than one half the Hausdorff dimension of the entire boundary Ω. We extend these results to BRW on free products of groups. We prove existence of the box-counting dimension, show that the Hausdorff dimension equals the box-counting dimension and present a formula in terms of generating functions of the underlying random walk, see Theorem 3.5. In the same way, we obtain a formula for the Hausdorff dimension of the whole space of ends, see Theorem 3.8. This eventually leads to the result that the Hausdorff dimension of Λ is not larger than one half the Hausdorff dimension of the entire boundary. Another consequence of the formula of the Hausdorff dimension is that the dimension varies continuously in the subcritical regime, see Theorem 3.10. This affirms the conjecture made in [1] for general non-amenable groups that the Hausdorff dimension of the limit set is continuous for $\lambda \neq \lambda_c$ and discontinuous at λ_c. As pointed out in [13], the very same phenomenon holds for other growth processes (for example, hyperbolic branching Brownian motion, isotropic contact process on homogeneous trees) that exhibit a phase transition between *weak* and *strong survival*.

In [13], the behaviour of the critical BRW on the free group was studied in more detail and two phenomena were observed. First, $\Phi(R) = \mathrm{HD}(\Omega)/2$ if and only if the underlying random walk is the simple random walk. This statement is not true for our more general setting since there are non-simple random walks that attain the maximal Hausdorff dimension $\mathrm{HD}(\Omega)/2$, see Remark 3.12 together with Example 3.14. Second, it was shown in [13] that $\Phi(R) - \Phi(\lambda) \sim C\sqrt{R-\lambda}$ as $\lambda \uparrow R$. For free products of groups this behaviour turns out to be more subtle: $\Phi(R) - \Phi(\lambda)$ may behave like $C(R-\lambda)$ or $C\sqrt{R-\lambda}$ depending on whether the Green function is differentiable at its radius of convergence or not.

The very same phenomena were also studied in the continuous setting. Lalley and Sellke [18] studied the phase transition for branching Brownian motion on the hyperbolic disc and

Karpelevich, Pechersky, and Suhov [14] generalized these results to higher-dimensional Lobachevsky spaces. Grigor'yan and Kelbert [11] studied recurrence and transience for branching diffusion processes on Riemannian manifolds. In Cammarota and Orsingher [3], first results on a 'linear' growing system of particles on the hyperbolic disc are given.

In the case of free products of groups $\Gamma = \Gamma_1 * \cdots * \Gamma_r$, where at least one of the factors is infinite, another phase transition occurs. The boundary Ω can be decomposed into up to $r+1$ direct summands. For $1 \leqslant i \leqslant r$, let Ω_i denote the set of ends described by semi-infinite non-backtracking paths, which eventually stay in one copy of Γ_i. The set Ω_∞ consists of all ends described by infinite, non-backtracking paths that change the different copies of the free factors infinitely many times. Now, for all infinite Γ_i, Theorem 3.1 gives a criterion whether $\Lambda \cap \Omega_i \neq \emptyset$ almost surely. In particular, it states that there exists a critical value λ_i such that $\lambda \leqslant \lambda_i$ is equivalent to $\Lambda \cap \Omega_i = \emptyset$ almost surely. In other words, if we increase the growth parameter λ, then more and more different parts of the boundary appear in Λ. However, even if $\Lambda \cap \Omega_i \neq \emptyset$, only the infinite words contribute to the Hausdorff dimension of Λ, see Corollary 3.7.

Finally, for the case of free products of *finite* groups we slightly adapt the metric defined on the boundary and obtain (following analogously the reasoning in [13]) a simpler formula for the Hausdorff dimension of Λ, see Corollary 3.16. Analogously, we obtain a formula for the Hausdorff dimension of Λ, if we have a BRW on free products by amalgamation of *finite* groups, see Corollary 3.18. In both cases, the Hausdorff dimension can be expressed through a Perron–Frobenius eigenvalue.

Let us remark that free products have been studied in great variety. Asymptotic behaviour of return probabilities of random walks on free products has been studied in many ways; e.g. Gerl and Woess [7, 24], Sawyer [22], Cartwright and Soardi [5], Lalley [15] and Candellero and Gilch [4]. For free products of finite groups, Mairesse and Mathéus [19] computed an explicit formula for the drift and asymptotic entropy. Gilch [9, 10] computed different formulas for the drift and also for the entropy for random walks on free products of graphs. Our proofs involve, in a very crucial way, generating function techniques for free products. These techniques were introduced independently and simultaneously by Cartwright and Soardi [5], Woess [24], Voiculescu [23] and McLaughlin [20]. In particular, we show that the Hausdorff dimension can be computed as the solution of a functional equation in terms of double generating functions.

The structure of the paper is as follows. In Section 2, we give an introduction to random walks on free products, generating functions, and branching random walks. In Section 3, we state our results and illustrate them with sample computations. The proofs are given in Section 4.

2. *Branching random walks on free products*

2.1. *Free products of groups and random walks*

Let $\mathcal{I} = \{1, 2, \ldots, r\}$ be a finite index set. Suppose we are given finitely generated groups Γ_i, $i \in \mathcal{I}$, where each Γ_i is generated by a symmetric generating set S_i (that is, $s \in S_i$ implies $s^{-1} \in S_i$) with identity e_i. Let $\Gamma_i^\times := \Gamma_i \setminus \{e_i\}$, for every $i \in \mathcal{I}$ and let $\Gamma_*^\times := \bigcup_{i \in \mathcal{I}} \Gamma_i^\times$. The *free product* $\Gamma := \Gamma_1 * \cdots * \Gamma_r$ is defined as the set

$$\{x_1 x_2 \ldots x_n \mid n \in \mathbb{N}, x_j \in \Gamma_*^\times, x_j \in \Gamma_k^\times \Rightarrow x_{j+1} \notin \Gamma_k^\times\} \cup \{e\}. \tag{2.1}$$

That is, each element of Γ is a *word* $x_1 \ldots x_n$ such that each letter (also called *block*) x_i is a non-trivial element of one of the factors and two consecutive letters are not from the same free factor Γ_i; e denotes the empty word. We exclude the trivial cases where Γ_i is the trivial group and the case $r = 2 = |\Gamma_1| = |\Gamma_2|$; see beginning of Section 2.2 for further remarks. The group

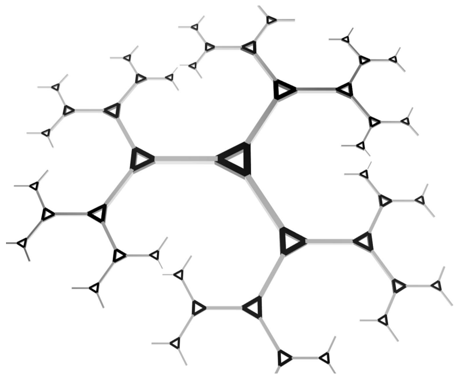


FIGURE 1. *Structure of the free product* $(\mathbb{Z}/2\mathbb{Z}) * (\mathbb{Z}/3\mathbb{Z})$.

operation on the free product Γ can be described as follows: if $u = u_1 \dots u_m, v = v_1 \dots v_n \in \Gamma$, then uv stands for their concatenation as words with possible contractions and cancellations in the middle in order to get the form of (2.1). For instance, if $u = aba$ and $v = abc$ with $a, c \in \Gamma_1^\times, b \in \Gamma_2^\times$ and $a^2 = e_1, b^2 \neq e_2$, then $uv = (aba)(abc) = a(b^2)c$. In particular, we set $ue_i := u$, for all $i \in \mathcal{I}$, and $eu := u$. Note that $\Gamma_i \subseteq \Gamma$ and e_i as a word in Γ is identified with e. The *block length* of a word $u = u_1 \dots u_m \in \Gamma$ is given by $\|u\| := m$. Additionally, we set $\|e\| := 0$. The *type* $\tau(u)$ of u is defined to be i if $u_m \in \Gamma_i^\times$; we set $\tau(e) := 0$.

To help visualizing the structure of a free product we may interpret the set Γ as the vertex set of its Cayley graph $\mathcal{X}$ (with respect to the generating set $\bigcup_{i\in\mathcal{I}} S_i$), which is constructed as follows: consider Cayley graphs $\mathcal{X}_1, \dots, \mathcal{X}_r$ of $\Gamma_1, \dots, \Gamma_r$ with respect to the (finite) symmetric generating sets $S_1, \dots, S_r$; take copies of $\mathcal{X}_1, \dots, \mathcal{X}_r$ and glue them together at their identities to one single common vertex, which becomes e; inductively, at each vertex $v = v_1 \dots v_k$ with $v_k \in \Gamma_i$ attach a copy of every $\mathcal{X}_j$, $j \neq i$, where v is identified with e_j of the new copy of $\mathcal{X}_j$ (see Figure 1). The natural graph distance on $\mathcal{X}$ is also used for elements of Γ and we write $l(u)$ for the *graph distance* or *length* of $u \in \Gamma$ to e. A *geodesic* of u is a shortest path from e to u. We remark that the length of an element may differ drastically from its block length.

We construct in a natural way a random walk on Γ from some given random walks on its free factors. Suppose we are given (symmetric, finitely supported) probability measures μ_i on Γ_i with $\langle \mathrm{supp}(\mu_i)\rangle = \Gamma_i$ for each $i \in \mathcal{I}$. For $x, y \in \Gamma_i$, the corresponding single-step transition probabilities of a random walk on Γ_i are given by $p_i(x, y) := \mu_i(x^{-1}y)$ and the n-step transition probabilities are denoted by $p_i^{(n)}(x, y) := \mu_i^{(n)}(x^{-1}y)$, where $\mu_i^{(n)}$ is the nth convolution power of μ_i. Each of these random walks is irreducible. For the sake of simplicity, we also assume $\mu_i(e_i) = 0$, for every $i \in \mathcal{I}$. We lift μ_i to a probability measure $\bar{\mu}_i$ on Γ by defining $\bar{\mu}_i(x) := \mu_i(x)$, if $x \in \Gamma_i$, and $\bar{\mu}_i(x) := 0$, otherwise. Let $\alpha_i > 0$, $i \in \mathcal{I}$, with $\sum_{i\in\mathcal{I}} \alpha_i = 1$. We now obtain a new finitely supported probability measure on Γ given by

$$\mu = \sum_{i\in\mathcal{I}} \alpha_i \bar{\mu}_i.$$

The random walk on Γ starting at e, which is governed by μ, is described by the sequence of random variables $(X_n)_{n\in\mathbb{N}_0}$. For $x, y \in \Gamma$, the associated single and n-step transition probabilities are denoted by $p(x, y) := \mu(x^{-1}y)$ and $p^{(n)}(x, y) := \mu^{(n)}(x^{-1}y)$, where $\mu^{(n)}$ is the nth convolution power of μ. The Cayley graph under consideration will always be with respect to the set of generators $\mathrm{supp}(\mu) = \bigcup_{i\in\mathcal{I}} \mathrm{supp}(\mu_i)$. We refer to Remark 3.11 for a short discussion for the case of non-nearest-neighbour random walks.

2.2. *Generating functions*

One key ingredient of the proofs is the study of the following generating functions. The most common among these generating functions are the *Green functions* related to μ_i and μ which are defined by

$$G_i(x_i,y_i|z) := \sum_{n\geqslant 0} p_i^{(n)}(x_i,y_i)\,z^n \quad \text{and} \quad G(x,y|z) := \sum_{n\geqslant 0} p^{(n)}(x,y)\,z^n,$$

where $z\in\mathbb{C}$, $i\in\mathcal{I}$, $x_i,y_i\in\Gamma_i$ and $x,y\in\Gamma$. We note that the free product Γ is non-amenable and that the radius of convergence R of $G(\cdot,\cdot|z)$ is strictly larger than 1; see, for example, [**25**, Theorem 10.10, Corollary 12.5]. In particular, this implies *transience* of our random walk on Γ. At this point, let us remark that the case $r=2=|\Gamma_1|=|\Gamma_2|$ leads to a recurrent random walk (and therefore to a recurrent BRW), which is the reason why we excluded this case. Moreover, non-amenability of Γ yields $G(e,e|R)<\infty$; see, for example, [**16**, Proposition 2.1].

The *first visit generating functions* related to μ_i and μ are given by

$$F_i(x_i,y_i|z) := \sum_{n\geqslant 0} \mathbb{P}[Y_n^{(i)}=y_i, \forall m\leqslant n-1: Y_m^{(i)}\neq y_i \mid Y_0^{(i)}=x_i]z^n$$

and

$$F(x,y|z) := \sum_{n\geqslant 0} \mathbb{P}[X_n=y, \forall m\leqslant n-1: X_m\neq y \mid X_0=x]z^n,$$

where $(Y_n^{(i)})_{n\in\mathbb{N}_0}$ describes a random walk on Γ_i governed by μ_i. For $M\subseteq\Gamma$, we also define

$$F(x,M|z) := \sum_{n\geqslant 0} \mathbb{P}[X_n\in M, \forall m\leqslant n-1: X_m\notin M \mid X_0=x]z^n$$

and the *first return generating function*

$$U(x,M|z) := \sum_{n\geqslant 1} \mathbb{P}[X_n\in M, \forall\, 1\leqslant m\leqslant n-1: X_m\notin M \mid X_0=x]z^n.$$

By a Harnack-type inequality, the generating functions $F(\cdot,\cdot|z)$ and $U(\cdot,\cdot|z)$ have also radii of convergence of at least $R>1$ and $U(x,M|z)=F(x,M|z)$, if $x\notin M$. By transitivity, we have $G_i(x_i,x_i|z)=G_i(e_i,e_i|z)$ and $G(x,x|z)=G(e,e|z)$, for all $x_i\in\Gamma_i$ and $x\in\Gamma$. For $x\in\Gamma\setminus\{e\}$, we have

$$G(e,e|z) > F(e,x|z)G(x,x|z)F(x,e|z); \tag{2.2}$$

indeed, while on the left-hand side we take into account all paths from e to e, on the right-hand side we only take into account all random walk paths from e to e which pass through x; therefore, strict inequality follows from irreducibility of the random walk which ensures always existence of random walk paths from e to e not passing through x. Symmetry of the laws μ_i now implies that $F(e,x|z)<1$ for all $|z|\leqslant R$ and all $x\in\Gamma\setminus\{e\}$. The *last visit generating functions* related to μ_i and μ are given by

$$L_i(x_i,y_i|z) := \sum_{n\geqslant 0} \mathbb{P}[Y_n^{(i)}=y_i, \forall 1\leqslant m\leqslant n: Y_m^{(i)}\neq x_i \mid Y_0^{(i)}=x_i]z^n$$

and

$$L(x,y|z) := \sum_{n\geqslant 0} \mathbb{P}[X_n=y, \forall 1\leqslant m\leqslant n: X_m\neq x \mid X_0=x]z^n.$$

We have the following important equations, which follow by conditioning on the first visits of y_i and y, the last visits of x_i and x, respectively:

$$\begin{aligned} G_i(x_i,y_i|z) &= F_i(x_i,y_i|z)\cdot G_i(y_i,y_i|z) = G_i(x_i,x_i|z)\cdot L_i(x_i,y_i|z),\\ G(x,y|z) &= F(x,y|z)\cdot G(y,y|z) = G(x,x|z)\cdot L(x,y|z). \end{aligned} \tag{2.3}$$

Thus, by transitivity, we obtain

$$F(x,y|z) = L(x,y|z) \quad \text{for any } x,y \in \Gamma \text{ and } |z| \leqslant R. \tag{2.4}$$

Let $x,y,w \in \Gamma$ such that all (random walk) paths from x to w pass through y. Then

$$F(x,w|z) = F(x,y|z) \cdot F(y,w|z) \quad \text{and} \quad L(x,w|z) = L(x,y|z) \cdot L(y,w|z); \tag{2.5}$$

this can be checked by conditioning on the first/last visit of y when walking from x to w. For $i \in \mathcal{I}$ and $z \in \mathbb{C}$, we define the functions

$$\xi_i(z) := U(e, \operatorname{supp}(\mu_i)|z) = U(e, \Gamma_i^\times|z) = F(e, \operatorname{supp}(\mu_i)|z), \tag{2.6}$$

which also have radii of convergence of at least $R > 1$. We remark that $\xi_i(1) < 1$; see, for example, [**9**, Lemma 2.3]. Moreover, we have $F(x_i,y_i|z) = F_i(x_i,y_i|\xi_i(z))$ and $L(x_i,y_i|z) = L_i(x_i,y_i|\xi_i(z))$, for all $x_i,y_i \in \Gamma_i$; see [**25**, Proposition 9.18c; **9**, Lemma 2.2]. Thus, by conditioning on the number of visits of e before finally making a step from e to $\Gamma_i^\times$ we obtain the following formula:

$$\xi_i(z) = \frac{\alpha_i z}{1 - \sum_{j\in\mathcal{I}\setminus\{i\}} \sum_{s\in\Gamma_j} \alpha_j \mu_j(s) z F_j(s, e_j|\xi_j(z))}. \tag{2.7}$$

Finally, we define the following power series that will lead to a useful expression for the Hausdorff dimension. Let

$$\mathcal{F}(\lambda|z) := \sum_{x\in\Gamma} F(e,x|\lambda)\, z^{l(x)}, \tag{2.8}$$

and define for $i \in \mathcal{I}$:

$$\mathcal{F}_i^+(\lambda|z) := \sum_{x\in\Gamma_i^\times} F(e,x|\lambda)\, z^{l(x)} = \sum_{x\in\Gamma_i^\times} F_i(e_i,x|\xi_i(\lambda)) z^{l(x)}, \tag{2.9}$$

$$\mathcal{F}_i(\lambda|z) := \sum_{n\geqslant 1} \sum_{\substack{x=x_1\ldots x_n\in\Gamma:\\ x_1\in\Gamma_i^\times}} F(e,x|\lambda)\, z^{l(x)} = \mathcal{F}_i^+(\lambda|z)\left(1 + \sum_{j\in\mathcal{I}\setminus\{i\}} \mathcal{F}_j(\lambda|z)\right). \tag{2.10}$$

The latter functions satisfy the following relation:

$$\mathcal{F}(\lambda|z) = 1 + \sum_{i\in\mathcal{I}} \mathcal{F}_i(\lambda|z). \tag{2.11}$$

2.3. *Branching random walks*

In this section, we introduce discrete-time branching random walks on free products and recall some basic results.

There are two different main descriptions or constructions of a BRW. The first defines the process inductively as a *growing cloud* of particles moving in (discrete) time and space. The second, via tree-indexed random walks, uses the fact that the branching distribution does not depend on the space. For that reason one can separate branching and movement into two steps. First, one generates the whole genealogy of the process and then one maps the corresponding genealogical tree into the Cayley graph. In both cases, we need the following definition. A Galton–Watson process is characterized through an *offspring distribution* ν. This is a probability measure on $\mathbb{N} = \{0,1,2,3,\ldots\}$ with mean (or also called growth parameter) $\lambda = \sum_{k\geqslant 1} k\,\nu(k) \in (0,\infty)$. We assume that ν has *finite second moment*, that is, $\sum_{k\geqslant 1} k^2$ $\nu(k) < \infty$. Moreover, we exclude the cases where $\nu(0) > 0$ and $\nu(1) = 1$; this guarantees that the process survives almost surely and that the BRW is not reduced to a (non-branching) random walk.

The BRW on Γ is defined inductively: at time 0 we have one particle at e (if not mentioned otherwise). Between time n and $n+1$ the process performs two steps: branching and movement.

First, each particle, independently of all others and the previous history of the process, produces descendants according to ν and dies. Second, each of these descendants, independently of all others and the past, moves to a neighbour vertex in Γ according to μ. A particle located at some vertex $x \in \Gamma$ at time n has a unique direct ancestor at time $n-1$. Consequently, each particle has a unique finite sequence of ancestors, the family history, which traces back to the original starting particle at e. The sequence of the locations of its ancestors (chronologically ordered) gives a path from e to x, which we call the *trail* of the particle.

Sometimes it will be convenient to work with the interpretation of a BRW as a tree-indexed random walk, see [2]. Let $\mathcal{T}$ be a rooted infinite tree. The root is denoted by $\mathbf{r}$ and other vertices by v and let $|v|$ be the (graph) distance from v to the root $\mathbf{r}$. The random walk on Γ indexed by $\mathcal{T}$ is the collection of Γ-valued random variables $(S_v)_{v\in\mathcal{T}}$ defined as follows. Label the edges of $\mathcal{T}$ with i.i.d. random variables η_v with distribution μ; the random variable η_v is the label of the edge (v^-, v). Define $S_v = e \cdot \prod_{i=1}^{|v|} \eta_{v_i}$, where $\langle v_0 = \mathbf{r}, v_1, \ldots, v_n = v\rangle$ is the unique geodesic (also called ancestry of v) from $\mathbf{r}$ to v at level n. A tree-indexed random walk becomes a BRW if the underlying tree is a Galton–Watson tree induced by ν. We refer to $\mathcal{T}$ as the family tree and to $\mathcal{X}$ as the base graph of the BRW. Furthermore, a vertex $v \in \mathcal{T}$ is called a particle of the BRW and $\mathcal{T}_n$ denotes the vertices of $\mathcal{T}$ on level n or equivalently the particles in generation n.

A useful variation of the first description of a BRW is the *coloured BRW*, see [13]. This process behaves like a standard BRW where in addition each particle is either blue or red. In order to define this coloured version we choose a subset M of Γ that plays the role of a 'paint bucket'. We start the BRW with one blue particle at e. Blue particles located outside of M produce blue offspring. A blue particle that hits the paint bucket is frozen there and will be replaced by a red particle. The new red particle starts an ordinary (red-coloured) BRW. As a consequence, every red particle has exactly one 'frozen' ancestor in M.

We denote by $Z_\infty(M) \in \mathbb{N} \cup \{\infty\}$ the random number of frozen (blue) particles in M during the whole branching process. If $M = \{x\}$, then we just write $Z_\infty(x)$.

For ease of presentation, we will switch freely between the different definitions of a BRW; nevertheless, it will always be clear from the context which description we are using.

A BRW on a Cayley graph is called *recurrent* if each vertex is visited infinitely many times and *transient* if any finite subset is eventually free of particles. The recurrence/transience behaviour is well understood. In fact, we have the following classification in recurrence and transience, see [2] for the sub- and supercritical cases and [6] for the critical case. We also refer to [11] for the corresponding result in the continuous setting.

THEOREM 2.1. *The BRW is transient if and only if $\lambda \leqslant R$.*

Recall that in the language of [13] transience is equivalent to *weak survival* if $\lambda > 1$. For the rest of this paper, we will restrict our investigation to the case of transience or weak survival. Since in this case the process eventually vacates every finite subset of Γ almost surely the investigation of the *convergence* of the BRW to the geometric boundary is meaningful.

2.4. *Ends of graphs, box-counting dimension and Hausdorff dimension*

Let us first recall some basic notations on infinite graphs. Let $\mathcal{G}$ be an infinite, connected, locally finite graph with countable vertex set and root e. For ease of presentation, we will identify $\mathcal{G}$ or a subgraph with its vertex set. A *path* of length n in $\mathcal{G}$ is a finite sequence of vertices $[x_0, x_1, \ldots, x_n]$ such that there is an edge from x_{i-1} to x_i, for each $i \in \{1, \ldots, n\}$. Recall that a *geodesic* of a vertex $x \in \mathcal{G}$ is a shortest path from e to x in $\mathcal{G}$. A *ray* is a semi-infinite path $[e = x_0, x_1, x_2, \ldots]$, which does not backtrack, that is, $x_i \neq x_j$, if $i \neq j$. Two rays η_1 and

η_2 are *equivalent* if there is a third ray which shares infinitely many vertices with η_1 and η_2. An equivalence class of rays is called an *end*. The set of equivalence classes of rays is called the *end boundary* of $\mathcal{G}$, denoted by $\partial\mathcal{G}$. For further details, we refer to [**25**, Section 21].

In the case of free products we have different types of ends occurring in the Cayley graph $\mathcal{X}$ of Γ: ends arising from ends in one of the $\mathcal{X}_i$, and 'infinite words'. More precisely, denote by $\Omega_i^{(0)}$ the set of ends of $\mathcal{X}_i$. For $\omega_i \in \Omega_i^{(0)}$, let $\eta = [e_i, y_1, y_2, \ldots] \in \omega_i$ and let $x \in \Gamma$, where $[x_0, x_1, \ldots, x_n]$ is a geodesic from x_0 to $x = x_n$. Then, the ray $x\eta := [x_0, x_1, \ldots, x_n, x_n y_1, x_n y_2, \ldots]$ describes an end in Γ. The end described by $x\eta$ is denoted by $x\omega_i$. We set $\Omega_i := \{x\omega_i \mid x \in \Gamma, \omega_i \in \Omega_i^{(0)}\}$. Moreover, the set of infinite words is given by

$$\Omega_\infty = \{x_1 x_2 x_3 \ldots \in (\Gamma_*^\times)^{\mathbb{N}} \,|\, x_j \in \Gamma_k^\times \Rightarrow x_{j+1} \notin \Gamma_k^\times\}.$$

It is easy to see that the set Ω of ends of $\mathcal{X}$ can be decomposed in the following way:

$$\Omega = \Omega_\infty \uplus \Omega_1 \uplus \Omega_2 \uplus \ldots \uplus \Omega_r.$$

Observe that Ω_i is empty if and only if Γ_i is finite. Thus, if all groups Γ_i are finite, then $\Omega = \Omega_\infty$.

In order to measure the size of Ω, we define a metric on Ω. We say that an end $\omega_1 \in \Omega$ is contained in a subset of the graph if all representatives have all but finitely many vertices in this subset. Now, if we remove from $\mathcal{X}$ any finite vertex subset $F \subseteq \mathcal{X}$ (including the removal of edges to vertices in F), then there is exactly one connected component in the reduced graph $\mathcal{X} \setminus F$ containing the end ω_1. We call this component the ω_1-component and say that ω_1 ends up in this component. Denote by $B_m := \{x \in \Gamma \mid l(x) \leqslant m\}$ the ball centred at e with radius m; we also set $B_{-1} := \emptyset$. Let $\omega_2 \in \Omega$ be another end with $\omega_1 \neq \omega_2$. Obviously, there is some maximal $m \in \mathbb{N}_0$ such that ω_1 and ω_2 end up in the same connected component of $\mathcal{X} \setminus B_{m-1}$. We write $c(\omega_1, \omega_2)$ for this maximal integer m. We now define a metric on Ω by

$$d_\Omega(\omega_1, \omega_2) := \alpha^{c(\omega_1, \omega_2)},$$

where $\alpha \in (0,1)$ is arbitrary, but fixed. Additionally, we set $d_\Omega(\omega_1, \omega_1) := 0$. The ball $B(\omega, \varepsilon)$ centred at $\omega \in \Omega$ with radius $\varepsilon \geqslant 0$ is given by all ends $\hat\omega \in \Omega$ with $d_\Omega(\omega, \hat\omega) \leqslant \varepsilon$. In other words, if $\varepsilon = \alpha^m$, then $\hat\omega \in B(\omega, \varepsilon)$ if and only if ω and $\hat\omega$ end up in the same component of $\mathcal{X} \setminus B_{m-1}$.

A *cover* of a subset $\Omega' \subseteq \Omega$ is a finite or countable set of balls of the form $B(\omega, \varepsilon_\omega)$ with $\omega \in \Omega'$ and $\varepsilon_\omega > 0$ such that the union of these balls include Ω'. For any $\varepsilon > 0$, let $N_\varepsilon(\Omega')$ be the minimal number of balls of the form $B(\omega, \varepsilon_\omega)$ with $\omega \in \Omega'$ and $0 < \varepsilon_\omega \leqslant \varepsilon$, which cover Ω'. Apparently, $N_\varepsilon(\Omega')$ is bounded from above by the number of elements in Γ at graph distance $m = \lceil \log(\varepsilon)/\log(\alpha) \rceil$. The *lower* and *upper box-counting dimension* (also called *the Minkowski dimension*) of Ω' are defined as

$$\underline{\mathrm{BD}}(\Omega') := \liminf_{\varepsilon \downarrow 0} \frac{\log N_\varepsilon(\Omega')}{-\log \varepsilon} \quad \text{and} \quad \overline{\mathrm{BD}}(\Omega') := \limsup_{\varepsilon \downarrow 0} \frac{\log N_\varepsilon(\Omega')}{-\log \varepsilon}. \tag{2.12}$$

If both limits are equal, then the common value is called the *box-counting dimension* $\mathrm{BD}(\Omega')$ of Ω'.

Another well-known measure for the size of Ω' is given by the Hausdorff dimension. For $\delta > 0$, the *δ-dimensional Hausdorff measure* of Ω' is defined by

$$\mathcal{H}_\delta(\Omega') := \liminf_{\varepsilon \downarrow 0} \left\{ \sum_i \varepsilon_i^\delta \,\middle|\, \{B(\cdot, \varepsilon_i)\}_i \text{ is a cover of } \Omega' \text{ with } \varepsilon_i < \varepsilon \right\}.$$

Then the *Hausdorff dimension* of Ω' is defined as

$$\mathrm{HD}(\Omega') := \inf\{\delta \geqslant 0 \,|\, \mathcal{H}_\delta(\Omega') = 0\}. \tag{2.13}$$

Since $\mathcal{X}$ has bounded vertex degrees we have $\mathrm{HD}(\Omega') < \infty$. It is well known that, for all $\Omega' \subseteq \Omega$,

$$\mathrm{HD}(\Omega') \leqslant \underline{\mathrm{BD}}(\Omega').$$

One of our main goals is to investigate to which kind of ends the BRW converges and to compare the dimensions of the whole space of ends with the set of ends which are 'hit' by the BRW. More precisely, for any $\omega \in \Omega$, if we remove any finite vertex subset $F \subseteq \mathcal{X}$, then there is exactly one connected component in the reduced graph $\mathcal{X} \setminus F$ containing ω. We say that the BRW *accumulates* at the end ω if for every finite vertex subset $F \subseteq \mathcal{X}$ there is at least one particle visiting the connected ω-component in $\mathcal{X} \setminus F$. The set of accumulation points is denoted by Λ. If the BRW is recurrent, then $\Omega = \Lambda$; thus, we restrict our investigation to the more interesting case of transience and therefore assume $1 < \lambda \leqslant R$. Note that $\Lambda \cap \Omega_\infty$ is almost surely non-empty; each infinite ancestry line converges to some element in $g_1 g_2 \ldots \in \Omega_\infty$ with convergence in the sense that the length of the common prefix of the particle's location and $g_1 g_2 \ldots$ tends to infinity, see, for example, [**9**, Proposition 2.5]. We also remark that the Hausdorff dimensions of Λ and $\Lambda \cap \Omega_\infty$ are almost surely constant, which can be shown analogously as explained in [**13**, Section 1, Remark (C)].

3. *Results*

In this section, we summarize our results about branching random walks on free products and present several explicit examples.

3.1. *Main results*

The first result describes how the structure of Λ gets richer when increasing the growth parameter λ and that there are up to $r = |\mathcal{I}|$ possible phase transitions.

THEOREM 3.1. *Let $\lambda \in (1, R]$. Then $\mathbb{P}[\Lambda \cap \Omega_i \neq \emptyset] \in \{0, 1\}$, and $\mathbb{P}[\Lambda \cap \Omega_i \neq \emptyset] = 1$ if and only if $\xi_i(\lambda) > 1$. More precisely*:

(1) *if $\xi_i(\lambda) \leqslant 1$, then $\emptyset \subsetneq \Lambda \subseteq \Omega_\infty$,*
(2) *if $\xi_i(\lambda) > 1$, then $\emptyset \subsetneq \Omega_\infty \cap \Lambda \subset \Lambda$ with $\Lambda \cap \Omega_i \neq \emptyset$ and $|\Lambda \cap \Omega_i| = \infty$.*

REMARK 3.2. In the case where one of the free factors is an infinite amenable group its ends do not appear in Λ. In other words, if $R_i = 1$ is the radius of convergence of $G_i(e_i, e_i|z)$, then $\xi_i(\lambda) \leqslant 1$ for all $\lambda \in (1, R]$; see [**25**, Lemma 17.1a]. Consequently, no ends in Ω_i contribute to Λ, that is, $\Lambda \cap \Omega_i = \emptyset$ almost surely.

We illustrate the above described behaviour in the following two examples:

EXAMPLE 3.3. Consider $\Gamma = \mathbb{Z}^{d_1} * \mathbb{Z}^{d_2}$ and let μ_1 and μ_2 be two symmetric probability measures on $\mathbb{Z}^{d_1}$ and $\mathbb{Z}^{d_2}$. Due to Kesten's amenability criterion, we have $R_1 = R_2 = 1$. Consequently, $\Lambda \subseteq \Omega_\infty$ almost surely, for all $\lambda \leqslant R$.

EXAMPLE 3.4. Consider $\Gamma = \Gamma_1 * \Gamma_2$, where Γ_1 and Γ_2 are non-amenable groups, and let μ_i define a symmetric random walk on Γ_i for $i \in \{1, 2\}$. Due to the non-amenability, we have that $R_1, R_2 > 1$ and $G_i(e_i, e_i|R_i) < \infty$. In the case where

$$\alpha_1 = \frac{R_1\, G_1(e_1, e_1|R_1)}{R_1\, G_1(e_1, e_1|R_1) + R_2\, G_2(e_2, e_2|R_2)},$$

we obtain by Woess [**25**, Lemma 17.1] that $\xi_1(R), \xi_2(R) > 1$. Therefore, there are numbers $\lambda_1, \lambda_2 \in (1, R)$ with $\xi_1(\lambda_1) = \xi_2(\lambda_2) = 1$ which leads to phase transitions at λ_1 and λ_2.

Now we state our first main result.

THEOREM 3.5. *Suppose that ν has a finite second moment. Then the box-counting dimension of Λ, $\Lambda \cap \Omega_\infty$, respectively, exists and equals the Hausdorff dimension of Λ, $\Lambda \cap \Omega_\infty$, respectively. Furthermore*:

$$\mathrm{BD}(\Lambda) = \mathrm{BD}(\Lambda \cap \Omega_\infty) = \mathrm{HD}(\Lambda) = \mathrm{HD}(\Lambda \cap \Omega_\infty) = \frac{\log z^*}{\log \alpha},$$

where z^* is the smallest real positive number with

$$\sum_{i\in\mathcal{I}} \frac{\mathcal{F}_i^+(\lambda|z^*)}{1+\mathcal{F}_i^+(\lambda|z^*)} = 1. \tag{3.1}$$

REMARK 3.6. The proof of Theorem 3.5 directly applies to BRW on free products of *finite graphs* and a corresponding result holds verbatim; see, for example [**25**, Section 9.C], for a formal definition of general free products and random walks on them.

As a first consequence, we obtain that only infinite words contribute to the dimension of Λ.

COROLLARY 3.7. *For $i \in \mathcal{I}$, $\mathrm{HD}(\Lambda \cap \Omega_i) < \mathrm{HD}(\Lambda \cap \Omega_\infty)$.*

For $i \in \mathcal{I}$, $m \in \mathbb{N}$ and $z \in \mathbb{C}$, we define $S_i(m) := |\{x \in \Gamma_i \mid l(x) = m\}|$ and

$$\mathcal{S}_i^+(z) := \sum_{m\geqslant 1} S_i(m) z^m.$$

Analogously to Theorem 3.5, we can prove existence of the box-counting dimension of the whole boundary Ω and express the dimension as the solution of a functional equation.

THEOREM 3.8. *The box-counting dimensions of Ω and Ω_∞ exist and satisfy*

$$\mathrm{BD}(\Omega) = \mathrm{BD}(\Omega_\infty) = \mathrm{HD}(\Omega) = \mathrm{HD}(\Omega_\infty) = \frac{\log z^*_{\mathcal{S}}}{\log \alpha},$$

where $z^*_{\mathcal{S}}$ is the smallest real positive number with

$$\sum_{i\in\mathcal{I}} \frac{\mathcal{S}_i^+(z^*_{\mathcal{S}})}{1+\mathcal{S}_i^+(z^*_{\mathcal{S}})} = 1. \tag{3.2}$$

Analogously to Corollary 3.7, we obtain that the Hausdorff dimension of Ω arises only from the ends in Ω_∞.

COROLLARY 3.9. *For all $i \in \mathcal{I}$, $\mathrm{HD}(\Omega_i) < \mathrm{HD}(\Omega_\infty)$.*

Beyond these first consequences of Theorems 3.5 and 3.8, the expressions for the Hausdorff dimensions allow us to study first regularity properties. For any fixed free product Γ, let us consider the function

$$\Phi : [1,\infty) \longrightarrow \mathbb{R} : \lambda \longmapsto \mathrm{HD}(\Lambda),$$

which assigns to every value λ the Hausdorff dimension of Λ of a BRW with growth parameter λ. The limit case $\lambda = 1$ corresponds to the degenerate case of a non-BRW; in this case, the Hausdorff dimension is just zero.

THEOREM 3.10. *The function $\Phi(\lambda)$ has the following properties*:

(1) $\Phi(\lambda)$ *is strictly increasing on* $[1, R]$, $\Phi(1) = 0$ *and* $\Phi(\lambda) = \mathrm{HD}(\Omega)$ *for all* $\lambda > R$;
(2) $\Phi(\lambda)$ *is continuous in* $[1,\infty) \setminus \{R\}$ *and continuous from the left at* $\lambda = R$ *with*

$$\Phi(R) \leqslant \tfrac{1}{2}\mathrm{HD}(\Omega),$$

(3) $\Phi(\lambda)$ *has the following behaviour as* $\lambda \uparrow R$:

$$\Phi(R) - \Phi(\lambda) \sim \begin{cases} C_1 \cdot (R-\lambda) & \text{if } G'(R) < \infty, \\ C_2 \cdot \sqrt{R-\lambda} & \text{if } G'(R) = \infty \end{cases}$$

for suitable constant C_1 and C_2, respectively.

REMARK 3.11. The last theorem states that $\mathrm{HD}(\Lambda)$ does not exceed $\mathrm{HD}(\Omega)/2$ unless the BRW is recurrent. We always assumed the random walk to be of nearest-neighbour type. However, we feel confident that our techniques work well in the case of finite range random walks and that the equality $\mathrm{HD}(\Lambda) \leqslant \mathrm{HD}(\Omega)/2$ does not depend on the choice of the metric. This type of phenomenon was already conjectured for the contact process on the homogeneous tree in [**17**]. We also refer to Section 8 in [**18**] for a discussion how the value $\frac{1}{2}$ can be explained through the 'backscattering principle'.

REMARK 3.12. In [**13**], it was shown that $\mathrm{HD}(\Lambda) = \mathrm{HD}(\Omega)/2$ only if $\lambda = R$ and if the underlying walk is a simple random walk. In our more general setting this is no longer true, since the maximal Hausdorff dimension can also be attained by a non-simple random walk, see Example 3.14. More generally, we conjecture that one has maximal dimension for the BRW (with λ being the critical growth value). for every choice of $\alpha_1 \in (0, 1)$. if we consider a general free product $\Gamma = \Gamma_1 * \Gamma_2$ with μ_1 and μ_2 governing positive recurrent random walks on the single factors Γ_1 and Γ_2.

REMARK 3.13. Recall that we always assume that the random walk on Γ is symmetric. This assumption can be dropped for free products of *finite* groups/graphs. In this case, we always have the crucial property $F(e, x|R) < 1$ for all $x \in \Gamma \setminus \{e\}$ (compare with (2.2)). In fact, if $x = x_1 \dots x_m \in \Gamma \setminus \{e\}$, then

$$F(e, x_1 \dots x_m|R) = \prod_{j=1}^{m} F_{\tau(x_j)}(e_{\tau(x_j)}, x_j \mid \xi_{\tau(x_j)}(R)) < 1,$$

as $\xi_i(R) < 1$ due to Woess [**25**, Lemma 17.1, Theorem 9.22].

Theorem 3.5 allows explicit calculations in all cases where formulas for the involved generating functions are known. In the following examples, we set the exponent of the metric on Ω equal to $\frac{1}{2}$, that is, $d_\Omega(\cdot,\cdot) = 2^{-c(\cdot,\cdot)}$.

EXAMPLE 3.14. Consider the free product $\Gamma = \Gamma_1 * \Gamma_2 = (\mathbb{Z}/3\mathbb{Z}) * (\mathbb{Z}/2\mathbb{Z})$, where $\mathbb{Z}/3\mathbb{Z} = \{e_1, a, a^2\}$, with $\mathrm{supp}(\mu_1) = \{a, a^2\}$. The required generating functions $F(e, x|\lambda)$, $x \in \Gamma_*^\times$, may, for example, be obtained by solving the finite systems of equations given in [**24**, Proposition 3c],

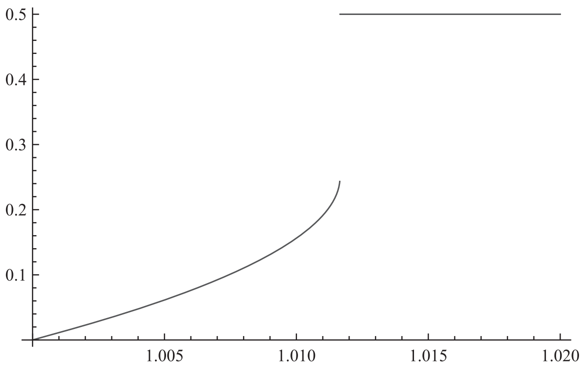


FIGURE 2. *Hausdorff dimension* HD(Λ) *of a BRW on* $(\mathbb{Z}/3\mathbb{Z}) * (\mathbb{Z}/2\mathbb{Z})$ *in dependence of* λ *on the* x*-axis.*

and therefore HD(Λ) can be computed via Equation (3.1). Solving Equation (3.2) leads to $\mathrm{HD}(\Omega) = \frac{1}{2}$. Figure 2 shows (with the help of numerical computation by MATHEMATICA) the graph of the function $\lambda \mapsto \mathrm{HD}(\Lambda)$ for simple random walk on $(\mathbb{Z}/3\mathbb{Z}) * (\mathbb{Z}/2\mathbb{Z})$. Let us remark that in this case the critical parameter R can be explicitly calculated by the formula given in [**25**, (9.29),(3)].

Another interesting phenomenon occurs in this example. If $\mu_1(a) = \mu_1(a^2) = \frac{1}{2}$ and if we let α_1 vary in the interval $(0,1)$ and denote by $R(\alpha_1)$ the radius of convergence of $G(e,e|z)$ in dependence of α_1, then we always obtain $\Phi(R(\alpha_1)) = \frac{1}{2}\mathrm{HD}(\Omega)$, which can be verified by explicit calculations with the help of MATHEMATICA.

EXAMPLE 3.15. We consider the free product of two infinite 'ladders' $\mathbb{Z} \times (\mathbb{Z}/2\mathbb{Z})$. We set $\alpha_1 = \alpha_2 = \frac{1}{2}$ and $\mu_1((\pm 1,0)) = \mu_1((0,1)) = \mu_2((\pm 1,0)) = \mu_2((0,1)) = \frac{1}{3}$. The functions $F_1((0,0),(z,a)|z)$ with $(z,a) \in \mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ can be computed by solving a system of equations as is shown in [**10**, Section 7.2]. In order to compute the Hausdorff dimension of Λ one has to solve, analogously to [**9**, Section 6.2]:

$$\frac{\lambda}{2}\frac{\xi_1(\lambda)}{\lambda - \xi_1(\lambda)} = \frac{\xi_1(\lambda)}{\begin{array}{c}1 - (2\xi_1(\lambda)/3)(F_1((0,0),(1,0)|\xi_1(\lambda)) + F_1((0,0),(-1,0)|\xi_1(\lambda)) \\ +F_1((0,0),(0,1)|\xi_1(\lambda)))\end{array}}.$$

In order to compute HD(Ω) we observe that $S_1(1) = 3$ and $S_1(m) = 4$, for $m \geqslant 2$. Hence, $\mathcal{S}_1^+(z) = \mathcal{S}_2^+(z) = 3z + 4z^2/(1-z)$. This yields $z^*_{\mathcal{S}} = \sqrt{5} - 2$. Numerical evaluations then lead to a picture qualitatively similar to Figure 2.

3.2. *Free products of finite groups*

In this subsection, we give a more explicit formula for the box-counting dimension with respect to a slightly changed metric on the boundary in the case of free products of *finite* groups. In this case, we have $\Omega = \Omega_\infty$. Throughout the whole subsection we do not need the assumption that the random walks on the factors are symmetric. For any $\omega_1 = x_1 x_2 \ldots, \omega_2 = y_1 y_2 \ldots \in \Omega_\infty$ with $\omega_1 \neq \omega_2$, we define the *confluent* $\omega_1 \wedge \omega_2$ of ω_1 and ω_2 as the word $x_1 \ldots x_k$ of maximal length with $x_i = y_i$, for all $1 \leqslant i \leqslant k$. If $x_1 \neq y_1$, then $\omega_1 \wedge \omega_2 := e$. The metric on the boundary Ω_∞ is defined by

$$d_\Omega^{\mathrm{fin}}(\omega_1,\omega_2) := \alpha^{\|\omega_1 \wedge \omega_2\|},$$

for any arbitrary but fixed $\alpha \in (0,1)$. With respect to this metric on Ω_∞, we can define analogously to (2.12) and (2.13) the upper box-counting dimension $\overline{\mathrm{BD}}^{\mathrm{fin}}(\Omega')$, the box-counting dimension $\mathrm{BD}^{\mathrm{fin}}(\Omega')$ and the Hausdorff dimension $\mathrm{HD}^{\mathrm{fin}}(\Omega')$, for any $\Omega' \subseteq \Omega_\infty$. We set $\mathcal{F}_i^+(\lambda) := \mathcal{F}_i^+(\lambda|1)$ and define the matrix $M = (m(i,j))_{i,j\in\mathcal{I}}$ by

$$m(i,j) := \begin{cases} \mathcal{F}_j^+(\lambda) & \text{if } i \neq j, \\ 0 & \text{if } i = j. \end{cases}$$

Since M is irreducible and has non-negative entries, the Perron–Frobenius eigenvalue exists and is denoted by θ.

Furthermore, define the matrix $D = (d(i,j))_{i,j\in\mathcal{I}}$ by $d(i,j) := |\Gamma_j| - 1$, if $i \neq j$, and $d_{i,i} := 0$, and denote by ϱ its Perron–Frobenius eigenvalue. With this notation we obtain:

Corollary 3.16.

$$\mathrm{BD}^{\mathrm{fin}}(\Lambda) = \mathrm{HD}^{\mathrm{fin}}(\Lambda) = -\frac{\log\theta}{\log\alpha} \quad \textit{and} \quad \mathrm{BD}^{\mathrm{fin}}(\Omega) = \mathrm{HD}^{\mathrm{fin}}(\Omega) = -\frac{\log\varrho}{\log\alpha}.$$

Let us remark that, in the case of $\Gamma = \Gamma_1 * \Gamma_2$ with $|\Gamma_1| = |\Gamma_2| < \infty$, we obtain the following explicit formulas for the dimensions:

$$\mathrm{BD}^{\mathrm{fin}}(\Lambda) = \mathrm{HD}^{\mathrm{fin}}(\Lambda) = -\frac{\log\sqrt{\mathcal{F}_1^+(\lambda)\mathcal{F}_2^+(\lambda)}}{\log\alpha}$$

and

$$\mathrm{BD}^{\mathrm{fin}}(\Omega) = \mathrm{HD}^{\mathrm{fin}}(\Omega) = -\frac{\log\sqrt{(|\Gamma_1|-1)(|\Gamma_2|-1)}}{\log\alpha}.$$

Example 3.17. Consider $\Gamma = (\mathbb{Z}/3\mathbb{Z}) * (\mathbb{Z}/2\mathbb{Z})$, where $\mathbb{Z}/3\mathbb{Z} = \{e_1, a, a^2\}$ and $\mathbb{Z}/2\mathbb{Z} = \{e_2, b\}$. We choose $\mu(a) = p \in (0.1, 0.7)$, $\mu(a^2) = q \in (0, 0.9-p)$ and $\mu(b) = 1-p-q$. We set $\alpha := \frac{1}{2}$ and $\lambda = 1.005$. Let us note that this choice of the parameters p and q lead to $R \geqslant 1.005$, which can be verified by numerical evaluation. For instance, in [**8**, Section 3.6.1] the required generating functions are computed. In Figure 3, we can see the behaviour of $\mathrm{HD}^{\mathrm{fin}}(\Lambda)$ with $\lambda = 1.005$ in dependence of the parameters p and q. The Hausdorff dimension of the whole space of ends is 0.5; compare with Example 3.14.

3.3. *Free products by amalgamation of finite groups*

An important generalization of free products are free products by amalgamation (of finite groups). Let $\Gamma_1, \ldots, \Gamma_r, H$, be finite groups such that each group Γ_i contains a subgroup H_i that is isomorphic to H. Let $\phi_i : H_i \to H$ be an isomorphism for each $i \in \{1, \ldots, r\}$. Moreover, let S_i be a generating set of Γ_i and R_i its relations. The *free product by amalgamation* with respect to the subgroup H is defined by

$$\begin{aligned}\Gamma_H &:= \Gamma_1 *_H \Gamma_2 *_H \ldots *_H \Gamma_r \\ &:= \langle S_1, \ldots, S_r \mid R_1, \ldots, R_n, \phi_j^{-1}(\phi_i(a)) = a \ \forall a \in H_i \ \forall i,j \in \mathcal{I}\rangle.\end{aligned}$$

For $i \in \mathcal{I}$, the quotient Γ_i/H_i consists of all left co-sets of the form $x_iH_i = \{x_ih \mid h \in H_i\}$, where $x_i \in \Gamma_i$. We fix a set of representatives $\mathcal{R}_i := \{g_{i,1} = e_i, g_{i,2}, \ldots, g_{i,n_i}\}$ for the elements of Γ_i/H_i, that is, for each $y_i \in \Gamma_i$, there is a unique $g_{i,k} \in \mathcal{R}_i$ with $y_i \in g_{i,k}H_i$. We write $\hat{\tau}(x) = i$, if $x \in \mathcal{R}_i \setminus \{e_i\}$. The amalgam Γ_H consists of all finite words of the form

$$x_1x_2\ldots x_nh \tag{3.3}$$

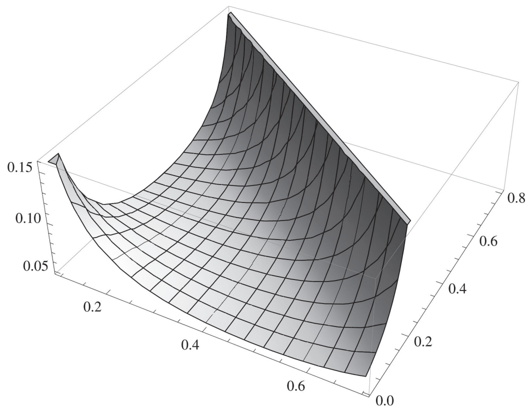


FIGURE 3. *Hausdorff dimension* HD(Λ) *of the BRW on* $(\mathbb{Z}/3\mathbb{Z}) * (\mathbb{Z}/2\mathbb{Z})$ *with* $\lambda = 1.005$ *in dependence of* p *and* q.

with $n \in \mathbb{N}_0$, $x_i \in \bigcup_{j\in\mathcal{I}} \mathcal{R}_j \setminus \{e_j\}$ and $h \in H$ such that $\hat\tau(x_i) \neq \hat\tau(x_{i+1})$. Here, without loss of generality, we may identify h with $\phi_1^{-1}(h)$, and e denotes again the empty word. Let Ω be the set of all ends of Γ_H, which consists of all infinite words of the form $w_1w_2\ldots \in (\bigcup_{i\in\mathcal{I}} \mathcal{R}_i \setminus \{e_i\})^{\mathbb{N}}$ such that $\hat\tau(w_i) \neq \hat\tau(w_{i+1})$, for all $i \in \mathbb{N}$. For any $\omega_1 = x_1x_2\ldots, \omega_2 = y_1y_2\ldots \in \Omega$ with $\omega_1 \neq \omega_2$, we define again the *confluent* $\omega_1 \wedge \omega_2$ of ω_1 and ω_2 as the word $x_1 \ldots x_k$ of maximal length with $x_i = y_i$, for all $1 \leqslant i \leqslant k$. If $x_1 \neq y_1$, then $\omega_1 \wedge \omega_2 := e$. Again we can define a metric on the boundary Ω:

$$d_\Omega^{(H)}(\omega_1, \omega_2) := \alpha^{\|\omega_1\wedge\omega_2\|},$$

for any $\alpha \in (0,1)$. With respect to this metric on Ω, we can define analogously to (2.12) and (2.13) the upper box-counting dimension $\overline{\mathrm{BD}}^{(H)}(\Omega')$, the box-counting dimension $\mathrm{BD}^{(H)}(\Omega')$ and the Hausdorff dimension $\mathrm{HD}^{(H)}(\Omega')$ for any $\Omega' \subseteq \Omega$.

Suppose we are given symmetric probability measures μ_i on the groups Γ_i and numbers $\alpha_i > 0$ such that $\sum_{i\in\mathcal{I}} \alpha_i = 1$. The random walk on Γ_H is then governed by

$$\mu(x) := \begin{cases} \alpha_i\mu_i(x) & \text{if } x \in \Gamma_i \setminus H_i, \\ \sum_{i\in\mathcal{I}} \alpha_i\mu_i(\phi_i^{-1}(\phi_1(x))) & \text{if } x \in H_1, \\ 0 & \text{otherwise.} \end{cases}$$

For $g_i \in \mathcal{R}_i$, denote by T_{g_iH} the stopping time of the first visit of the set g_iH_i. We introduce the following generating functions:

$$F_H(gh|z) := \sum_{n\geqslant 0} \mathbb{P}[T_{gH} = n, X_n = gh \mid X_0 = e]\, z^n,$$

where $g \in \bigcup_{i\in\mathcal{I}} \mathcal{R}_i \setminus \{e_i\}$, $h \in H_i$ and $z \in \mathbb{C}$. By symmetry we have $F_H(gh|z) \leqslant F(e, gh|z) < 1$; compare with (2.2). Conditioning on the first step of the random walk, we obtain

$$\begin{aligned} F_H(gh|z) = \mu(gh)z + \sum_{g_0\in\Gamma_{\tau(g)}\setminus gH_{\tau(g)}} \mu(g_0)zF_H(g_0^{-1}gh|z) \\ + \sum_{i\in\mathcal{I}\setminus\{\tau(g)\}} \sum_{g_0\in\Gamma_i} \mu(g_0)z \sum_{h_0\in H_i} F_H(g_0^{-1}h_0|z)F_H(h_0^{-1}gh|z). \end{aligned} \tag{3.4}$$

Since there are only *finitely* many functions $F_H(\cdot|z)$, one can compute $F_H(\cdot|z)$ by solving the finite system of quadratic equations (3.4). We also define

$$\mathcal{F}_i^{(H)}(z) := \sum_{\substack{g\in\mathcal{R}_i\setminus\{e_i\},\\ h\in H_i}} F_H(gh|z)$$

and the matrix $N = (n(i,j))_{i,j\in\mathcal{I}}$ with entries

$$n(i,j) := \begin{cases} \mathcal{F}_j^{(H)}(\lambda) & \text{if } i \neq j, \\ 0 & \text{if } i = j. \end{cases}$$

We denote by θ_H the Perron–Frobenius eigenvalue of N. Furthermore, we denote by ϱ_H the Perron–Frobenius eigenvalue of the matrix $D_H = (d_H(i,j))_{i,j\in\mathcal{I}}$, which is defined by

$$d_H(i,j) := \begin{cases} [\Gamma_j : H_j] - 1 & \text{if } i \neq j, \\ 0 & \text{if } i = j. \end{cases}$$

Finally, we can state the following formulas for the dimensions:

Corollary 3.18.

$$\mathrm{BD}^{(H)}(\Lambda) = \mathrm{HD}^{(H)}(\Lambda) = -\frac{\log\theta_H}{\log\alpha} \quad \textit{and} \quad \mathrm{BD}^{(H)}(\Omega) = \mathrm{HD}^{(H)}(\Omega) = -\frac{\log\varrho_H}{\log\alpha}.$$

Example 3.19. Consider the amalgam $(\mathbb{Z}/6\mathbb{Z}) *_{\mathbb{Z}/2\mathbb{Z}} (\mathbb{Z}/6\mathbb{Z})$. Hence, let $\Gamma_1 = \langle a \mid a^6 = e_1\rangle$, $\Gamma_2 = \langle b \mid b^6 = e_2\rangle$, and $H = \langle c \mid c^2 = e_H\rangle$, where e_H is the identity in H. The isomorphisms are defined through $\phi_1(a^3) = c = \phi_2(b^3)$. Eventually,

$$(\mathbb{Z}/6\mathbb{Z}) *_{\mathbb{Z}/2\mathbb{Z}} (\mathbb{Z}/6\mathbb{Z}) = \langle a, b \mid a^6 = b^6 = e, a^3 = b^3\rangle.$$

We set $\mu_1(a) = \mu_1(a^5) = \mu_2(b) = \mu_2(b^5) = \frac{1}{2}$, $\alpha_1 = \alpha_2 = \frac{1}{2}$ and consider the distance with base $\alpha = \frac{1}{2}$. The system (3.4) becomes then

$$F_H(a|z) = \frac{z}{4} + \frac{z}{4}F_H(a^2|z) + \frac{z}{2}\big(F_H(a|z)^2 + F_H(a^2|z)^2\big),$$
$$F_H(a^2|z) = \frac{z}{4}F_H(a|z) + \frac{z}{2}\big(F_H(a|z)F_H(a^2|z) + F_H(a^2|z)F_H(a|z)\big).$$

Observe that $F_H(a|z) = F_H(a^5|z)$ and $F_H(a^2|z) = F_H(a^4|z)$. The Hausdorff dimension of the BRW is then given by

$$\mathrm{HD}^{(H)}(\Lambda) = \frac{\log(2F_H(a|\lambda) + 2F_H(a^2|\lambda))}{\log 2},$$

while $\mathrm{HD}^{(H)}(\Omega) = 1$. The behaviour of $\mathrm{HD}^{(H)}(\Lambda)$ in function of λ is qualitatively the same as in Figure 2.

4. *Proofs*

4.1. *Proof of Theorem* 3.1

We first introduce some preliminary results on BRW. Using the description of a tree-indexed random walk it is easy to see that the distribution of the location of some particle in generation n has the same distribution as the location of a (non-branching) random walk on Γ after n steps, see [**2**].

LEMMA 4.1. *Let $v \in \mathcal{T}$ with $|v| = n$, for some $n \geqslant 1$. Then,*

$$\mathbb{P}[S_v = y] = P[X_n = y] = \mu^{(n)}(y).$$

The following lemma will be used several times in our proofs. It gives a formula for the expected number of elements frozen in a set M, in the coloured branching random walk. This observation can be found for example in [21] or [13, Lemma 1]. Nevertheless, we give a short proof since it is one of the essential points where the generating function $F(\cdot,\cdot|z)$ intervenes.

LEMMA 4.2. *For any $M \subseteq \Gamma$, we have $\mathbb{E}[Z_\infty(M)] = F(e, M|\lambda)$.*

Proof. For any $v \in \mathcal{T}$, let $\langle v_0 = \mathbf{r}, v_1, \ldots, v_{|v|} = v\rangle$ be the unique geodesic from $\mathbf{r}$ to v. Now, we define, for any $n \in \mathbb{N}$,

$$\mathrm{Fr}_v^{(n)} := \begin{cases} 1 & \text{if } v \in M \text{ and } v_i \notin M \quad \forall i \leqslant n-1, \\ 0 & \text{otherwise.} \end{cases}$$

In words, $\mathrm{Fr}_v^{(n)}$ is the number of particles being frozen in v at time n. Using the well-known fact that $E[|\mathcal{T}_n|] = \lambda^n$, we obtain

$$\begin{aligned} \mathbb{E}\left[\sum_{v \in \mathcal{T}_n} \mathrm{Fr}_v^{(n)}\right] &= \sum_{k \geqslant 1} \mathbb{E}\left[\sum_{v \in \mathcal{T}_n} \mathrm{Fr}_v^{(n)} \,\middle|\, |\mathcal{T}_n| = k\right] \mathbb{P}[|\mathcal{T}_n| = k] \\ &= \sum_{k \geqslant 1} \mathbb{P}[X_n \in M, \forall\, m \leqslant n-1 : X_m \notin M]\, k\, \mathbb{P}[|\mathcal{T}_n| = k] \\ &= \mathbb{P}[X_n \in M, \forall\, m \leqslant n-1 : X_m \notin M]\, \lambda^n. \end{aligned}$$

Summing over n finishes the proof. □

The proof of Theorem 3.1 splits up into the proofs of the following Propositions 4.3–4.5. Recall from the definition of $\Omega_i^{(0)}$ and Ω_i that $\Omega_i^{(0)} \subseteq \Omega_i \subseteq \Omega$.

PROPOSITION 4.3. *Ends of $\Omega_i^{(0)}$ occur in Λ with positive probability if and only if $\xi_i(\lambda) > 1$, that is, $\mathbb{P}[\Lambda \cap \Omega_i^{(0)} \neq \emptyset] > 0$ if and only if $\xi_i(\lambda) > 1$.*

Proof. It is convenient to work with the *coloured* BRW. In fact, the idea of the proof is to define an embedded Galton–Watson process that counts the number of particles that hit Γ_i, where $\xi_i(\lambda)$ will be the growth parameter.

We start the BRW with one particle in $e = e_i$. The first generation of the branching process is formed by those particles that are frozen in $\Gamma_i^\times$. Let us check that the number of those particles is almost surely finite. Since μ has finite support, every particle visiting $\Gamma_i^\times$ has to pass through $\mathrm{supp}(\mu_i)$. Hence, $Z_\infty(\Gamma_i^\times) = Z_\infty(\mathrm{supp}(\mu_i))$, which is almost surely finite since the BRW is transient. The second generation of the branching process is constructed as follows. For each particle frozen in some $x \in \Gamma_i^\times$, we start a new BRW where each particle when reaching $\Gamma_i \setminus \{x\}$ is frozen. Now, the second generation of the branching process consists of all these new frozen particles. Further generations are constructed inductively in the same way. Let ψ_n be the number of particles of this process at generation n. Obviously, $(\psi_n)_{n \geqslant 0}$ turns out to be a Galton–Watson process with mean

$$m_i = \mathbb{E}[Z_\infty(\mathrm{supp}(\mu_i))] = F(e, \mathrm{supp}(\mu_i)|\lambda) = \xi_i(\lambda).$$

Hence, this Galton–Watson process survives with positive probability if and only if $\xi_i(\lambda) > 1$; see, for example, [**12**, Theorem 6.1]. As a consequence, we have that Γ_i is visited infinitely many times with positive probability if $\xi_i(\lambda) > 1$. That is, $\mathbb{P}[\Lambda \cap \Omega_i^{(0)} \neq \emptyset] > 0$ if $\xi_i(\lambda) > 1$. On the other hand, $\xi_i(\lambda) \leqslant 1$ implies that Γ_i is almost surely visited only for a finite number of times and hence $\mathbb{P}[\Lambda \cap \Omega_i^{(0)} \neq \emptyset] = 0$. □

The next step is to show $|\Lambda \cap \Omega_i| = \infty$ if $\xi_i(\lambda) > 1$.

Proposition 4.4. *If $\xi_i(\lambda) > 1$, then there are almost surely infinitely many cosets $x\Gamma_i$, where the branching random walk accumulates. That is, the set*

$$\{x \in \Gamma \mid \tau(x) \neq i, x\Omega_i^{(0)} \cap \Lambda \neq \emptyset\}$$

is almost surely infinite.

Proof. We construct the family tree $\mathcal{T}$ of the BRW with branching distribution ν in the following way. We start with one geodesic line $v_\infty = \langle \mathbf{r}, v_1, v_2, \ldots \rangle$ and attach to each of the vertices independent copies of Galton–Watson trees where the distribution of the first generation is $\tilde{\nu}(k) = \nu(k+1)$ for $k \geqslant 0$ and ν for the other generations. The trajectory along v_∞ has the same distribution as a non-BRW, compare with Lemma 4.1. Hence, S_{v_n} converges almost surely to a random infinite word $g_\infty = g_1 g_2 \ldots \in \Omega_\infty$ as $n \to \infty$; here, we mean convergence in the sense that the block length of the common prefix of the location of S_{v_n} and g_∞ tends to infinity. Moreover, we define the random indices $n_1 := \min\{m \in \mathbb{N} \mid g_m \in \Gamma_i\}$, and recursively $n_k := \min\{m \in \mathbb{N} \mid m > n_{k-1}, g_m \in \Gamma_i\}$. Note that these indices are almost surely finite; see, for example, [**9**, Section 7.I]. Denote by $\hat{v}_k$ the first vertex in v_∞ with $\hat{v}_k = g_1 \ldots g_{n_k}$. Let B_k be the set of offspring of $\hat{v}_k = v_s$ different from v_{s+1} and denote by Λ_v the set of accumulation points of the descendants of some $v \in \mathcal{T}$. Moreover, we define A_k as the event that $\Lambda_v \cap S_v\Omega_i^{(0)} \neq \emptyset$ for some $v \in B_k$ with $\tau(v) = i$. Observe that the events A_k are i.i.d. since transitivity yields $\mathbb{P}[\Lambda_v \cap S_v\Omega_i^{(0)} \neq \emptyset] = \mathbb{P}[\Lambda \cap \Omega_i^{(0)} \neq \emptyset]$, for every $v \in \mathcal{T}$. Now, due to Proposition 4.3 and the fact that

$$\begin{aligned} P[B_k \neq \emptyset, \exists v \in B_k : \tau(S_v) = i] &= (1 - \nu(1)) \cdot \mathbb{P}[v \in B_k : \tau(S_v) = i \mid B_k \neq \emptyset] \\ &\geqslant (1 - \nu(1)) \cdot \alpha_i > 0, \end{aligned}$$

we have $P[A_k] \geqslant c$ for all k and some $c > 0$. Eventually, the Lemma of Borel–Cantelli yields that an infinite number of events A_k occurs almost surely. □

In order to complete the proof of Theorem 3.1 it remains to treat the critical and subcritical cases.

Proposition 4.5. *If $\xi_i(\lambda) \leqslant 1$, then $\mathbb{P}[\Lambda \cap \Omega_i \neq \emptyset] = 0$.*

Proof. Due to Proposition 4.3, we have that $\mathbb{P}[\Lambda \cap x\Omega_i^{(0)} \neq \emptyset] = 0$, for all $x \in \Gamma$: indeed, each $x \in \Gamma$ is almost surely visited finitely often; each particle, which hits x, starts its own BRW at x and each of these BRW hits $x\Omega_i^{(0)}$ only finitely often with probability 1. Since

$$\Lambda \cap \Omega_i = \biguplus_{x \in \Gamma : \tau(x) \neq i} (\Lambda \cap x\Omega_i^{(0)})$$

we conclude

$$\mathbb{P}[\Lambda \cap \Omega_i \neq \emptyset] = \sum_{x\in\Gamma:\tau(x)\neq i} \mathbb{P}[\Lambda \cap x\Omega_i^{(0)} \neq \emptyset] = 0. \qquad \square$$

4.2. *Proof of Theorem* 3.5 *and Corollary* 3.7

First, we show that the proposed formula for the dimension is an upper bound for the upper box-counting dimension; see Proposition 4.9 in Section 4.2.1. In the second step, we show that the proposed formula is also a lower bound for the Hausdorff dimension of Λ; see Corollary 4.14 in Subsection 4.2.2. Finally, this will imply the proof of Theorem 3.5 and Corollary 3.7.

4.2.1. *Upper bound for the box-counting dimension.* In this part, we show that $\log z^*/\log\alpha$ is an upper bound for $\overline{\mathrm{BD}}(\Lambda)$. To this end, we introduce the following notation: for $n\in\mathbb{N}$, we denote by

$$\mathcal{H}_n := \{x\in\Gamma \mid l(x)=n, x \text{ is visited by the BRW}\}$$

the set of visited sites at graph distance n. An important observation is that, for each end $\omega\in\Lambda$ and every $m\in\mathbb{N}$, the branching random walk has to visit at least one vertex $x_\omega\in\mathcal{H}_m$, where x_ω is in the ω-component of $\mathcal{X}\setminus B_{m-1}$. Thus,

$$\Lambda \subseteq \bigcup_{x\in\mathcal{H}_m} \{\omega\in\Omega \mid x \text{ lies in the } \omega\text{-component of } \mathcal{X}\setminus B_{m-1}\}.$$

This implies that Λ can be covered by $|\mathcal{H}_m|$ balls of radius α^m. Our strategy for the upper bound is to study the limit behaviour of $\mathbb{E}|\mathcal{H}_m|^{1/m}$ first and then the resulting limit behaviour of $|\mathcal{H}_m|^{1/m}$ as $m\to\infty$; see Lemma 4.8. This will eventually lead to the proposed upper bound for $\overline{\mathrm{BD}}(\Lambda)$; see Proposition 4.9.

Observe that $x\in\mathcal{H}_m$ if and only if $Z_\infty(x)\geqslant 1$. Therefore, by Lemma 4.2,

$$1\leqslant \mathbb{E}|\mathcal{H}_m| \leqslant \sum_{x\in\Gamma:l(x)=m} \mathbb{E}Z_\infty(x) = \sum_{x\in\Gamma:l(x)=m} F(e,x|\lambda) =: H_m.$$

We have that $H_{m+n}\leqslant H_m H_n$ and hence Fekete's lemma implies that $\lim_{n\to\infty} H_m^{1/m}$ exists. Recall the definitions of $\mathcal{F}(\lambda|z)=\sum_{m\geqslant 0} H_m\, z^m$, $\mathcal{F}_i^+(\lambda|z)$ and $\mathcal{F}_i(\lambda|z)$ in (2.8)–(2.10). Due to (2.11), we obtain the equation

$$\mathcal{F}_i(\lambda|z) = \mathcal{F}_i^+(\lambda|z)(\mathcal{F}(\lambda|z)-\mathcal{F}_i(\lambda|z)),$$

or equivalently

$$\mathcal{F}_i(\lambda|z) = \mathcal{F}(\lambda|z)\frac{\mathcal{F}_i^+(\lambda|z)}{1+\mathcal{F}_i^+(\lambda|z)}.$$

Hence,

$$\mathcal{F}(\lambda|z) = 1+\sum_{i\in\mathcal{I}}\mathcal{F}_i(\lambda|z) = 1+\mathcal{F}(\lambda|z)\sum_{i\in\mathcal{I}}\frac{\mathcal{F}_i^+(\lambda|z)}{1+\mathcal{F}_i^+(\lambda|z)},$$

or equivalently

$$\mathcal{F}(\lambda|z) = \frac{1}{1-\sum_{i\in\mathcal{I}}(\mathcal{F}_i^+(\lambda|z)/(1+\mathcal{F}_i^+(\lambda|z)))}. \tag{4.1}$$

This equation holds, for every $z\in\mathbb{C}$ with $|z|<R(\mathcal{F})$, where $R(\mathcal{F})$ is the radius of convergence of $\mathcal{F}(\lambda|z)$. Since

$$1\leqslant \lim_{m\to\infty} H_m^{1/m} = 1/R(\mathcal{F}),$$

we have

$$R(\mathcal{F})\leqslant 1. \tag{4.2}$$

In order to determine $R(\mathcal{F})$, we have to find (by Pringsheim's Theorem) the smallest singularity point on the positive x-axis of $\mathcal{F}(\lambda|z)$. This smallest singularity point is either one of the radii of convergence $R(\mathcal{F}_i^+)$ of the functions $\mathcal{F}_i^+(\lambda|z)$ or the smallest real positive number z^* with

$$\sum_{i\in\mathcal{I}} \frac{\mathcal{F}_i^+(\lambda|z^*)}{1+\mathcal{F}_i^+(\lambda|z^*)} = 1. \tag{4.3}$$

The next two lemmas imply that in fact $R(\mathcal{F}) = z^*$.

Lemma 4.6. *We have* $R(\mathcal{F}) \in (0,1)$.

Proof (This short proof was suggested by the referee). The fact that $R(\mathcal{F}) > 0$ follows from the fact that the Cayley graph grows not faster than exponentially. To see that $R(\mathcal{F}) < 1$ recall that Equation (2.3) states that the generating functions $F(e,x|z)$ and $G(e,x|z)$ are comparable, that is, $G(e,x|\lambda) = F(e,x|\lambda)G(e,e|\lambda)$. Hence, for some $C>0$, we have, for all $m\in\mathbb{N}$, that

$$\sum_{x:l(x)\leqslant m} F(e,x|\lambda) \geqslant C \sum_{x:l(x)\leqslant m} G(e,x|\lambda).$$

The sum on the right-hand side is the expected number of visits of the BRW in the ball B_m, the set of vertices $x\in\Gamma$ with $l(x)\leqslant m$. As we assumed the random walk to be of nearest-neighbour type all particles up to generation m must be contained in the ball B_m. The expected population size at time m is just λ^m which eventually implies that H_m grows exponentially fast, since $\lim_{m\to\infty} H_m^{1/m}$ exists and is at least 1. □

Lemma 4.7. *For all* $i\in\mathcal{I}$, $R(\mathcal{F}) = z^* < R(\mathcal{F}_i^+)$.

Proof. Let us first consider the case $\xi_i(\lambda) < 1$, where we obtain

$$\begin{aligned}\mathcal{F}_i^+(\lambda|1) &= \sum_{x\in\Gamma_i^\times} F_i(e_i,x|\xi_i(\lambda))\\ &= \frac{1}{G_i(e_i,e_i|\xi_i(\lambda))}\sum_{x\in\Gamma_i} G_i(e_i,x|\xi_i(\lambda)) - 1\\ &= \frac{1}{G_i(e_i,e_i|\xi_i(\lambda))(1-\xi_i(\lambda))} - 1 < \infty.\end{aligned} \tag{4.4}$$

Hence, $\xi_i(\lambda) < 1$ implies $R(\mathcal{F}_i^+) \geqslant 1 > R(\mathcal{F})$. In the case of $\xi_i(\lambda)\geqslant 1$, the claim follows from the following inequality:

$$\frac{1}{R(\mathcal{F})} = \limsup_{n\to\infty}\left(\sum_{\substack{x\in\Gamma:\\ l(x)=n}} F(e,x|\lambda)\right)^{1/n} > \limsup_{n\to\infty}\left(\sum_{\substack{x\in\Gamma_1:\\ l(x)=n}} F(e,x|\lambda)\right)^{1/n} = \frac{1}{R(\mathcal{F}_1^+)}. \tag{4.5}$$

In order to prove (4.5) we define, for $n\in\mathbb{N}$,

$$a_n := \log \sum_{\substack{x\in\Gamma_1:\\ l(x)=n}} F(e,x|\lambda).$$

We have that $a_n \geqslant 0$, since

$$\sum_{\substack{x\in\Gamma_1:\\ l(x)=n}} F(e,x|\lambda) = \sum_{\substack{x\in\Gamma_1:\\ l(x)=n}} F_1(e_1,x|\xi_1(\lambda)) \geqslant \sum_{\substack{x\in\Gamma_1:\\ l(x)=n}} F_1(e_1,x|1) \geqslant \mathbb{P}[T_{S_1(n)}<\infty] = 1,$$

where $S_1(n) := \{x \in \Gamma_1 \mid l(x) = n\}$ and T_M is the stopping time for the random walk on Γ_1 (governed by μ_1) of the first visit of a set $M \subseteq \Gamma_1$. Furthermore, $(a_n)_{n\in\mathbb{N}}$ is a subadditive sequence, that is, $a_m + a_n \geqslant a_{m+n}$ for all $m, n \in \mathbb{N}$. By Fekete's Lemma, the limit $\lim_{n\to\infty} a_n/n$ exists and is equal to $\inf_{n\in\mathbb{N}} a_n/n$, hence

$$\lim_{n\to\infty}\left(\sum_{\substack{x\in\Gamma_1:\\ l(x)=n}} F(e,x|\lambda)\right)^{1/n} = \frac{1}{R(\mathcal{F}_1^+)} = \inf_{n\in\mathbb{N}}\left(\sum_{\substack{x\in\Gamma_1:\\ l(x)=n}} F(e,x|\lambda)\right)^{1/n}.$$

The last equation implies that

$$\left(\sum_{\substack{x\in\Gamma_1:\\ l(x)=n}} F(e,x|\lambda)\right)^{1/n} \geqslant \frac{1}{R(\mathcal{F}_1^+)} =: q_1 \quad \forall n\in\mathbb{N}.$$

Observe that $\sum_{x\in\Gamma_2: l(x)=1} F(e,x|\lambda) \geqslant \xi_2(\lambda)$. Then, for all $n \in \mathbb{N}$:

$$\begin{aligned}
H_n = \sum_{\substack{x\in\Gamma:\\ l(x)=n}} F(e,x|\lambda) &= \sum_{k=1}^{n} \sum_{\substack{x=x_1\dots x_k\in\Gamma:\\ l(x)=n}} \prod_{j=1}^{k} F(e,x_j|\lambda)\\
&\geqslant \sum_{k=1}^{\lfloor n/2\rfloor} \sum_{\substack{x_1,\dots,x_k\in\Gamma_1:\\ l(x_1)+\dots l(x_k)+k=n}} \xi_2(\lambda)^k \prod_{j=1}^{k} F(e,x_j|\lambda)\\
&\geqslant \sum_{k=1}^{\lfloor n/2\rfloor} \sum_{\substack{n_1,\dots,n_k\in\mathbb{N}:\\ n_1+\dots+n_k+k=n}} q_1^{n_1}\xi_2(\lambda)q_1^{n_2}\xi_2(\lambda)q_1^{n_3}\dots\xi_2(\lambda)q_1^{n_k}\xi_2(\lambda)\\
&\geqslant \sum_{k=1}^{\lfloor n/2\rfloor} q_1^{n-k}\xi_2(\lambda)^k\binom{n-2k+k-1}{k-1}.
\end{aligned}$$

In the last inequality the binomial coefficients arise as follows: we think of counting the number of possibilities of placing $n-k$ (undistinguishable) balls into k urns, where each urn should at least contain one ball. We note that $n-k-2 \geqslant \lfloor n/2\rfloor - 1$, for all $k \leqslant \lfloor n/2\rfloor - 1$. Therefore, with the help of the Binomial theorem we obtain:

$$\begin{aligned}
H_n &\geqslant q_1^n \sum_{k=0}^{\lfloor n/2\rfloor-1}\left(\frac{\xi_2(\lambda)}{q_1}\right)^{k+1}\binom{n-k-2}{k}\\
&\geqslant q_1^{n-1}\xi_2(\lambda)\sum_{k=0}^{\lfloor n/2\rfloor-1}\left(\frac{\xi_2(\lambda)}{q_1}\right)^{k}\binom{\lfloor n/2\rfloor-1}{k} \geqslant q_1^{n-1}\xi_2(\lambda)\left(1+\frac{\xi_2(\lambda)}{q_1}\right)^{\lfloor n/2\rfloor-1}.
\end{aligned}$$

Taking nth roots on both sides and letting $n\to\infty$ yields

$$\liminf_{n\to\infty}\left(\sum_{\substack{x\in\Gamma:\\ l(x)=n}} F(e,x|\lambda)\right)^{1/n} \geqslant \frac{1}{R(\mathcal{F}_1^+)}\sqrt{1+\frac{\xi_2(\lambda)}{q_1}} > \frac{1}{R(\mathcal{F}_1^+)}. \tag{4.6}$$

□

The next lemma gives an almost sure upper bound for $|\mathcal{H}_m|^{1/m}$ as $m\to\infty$. Its proof is a straightforward application of Markov's Inequality and the Lemma of Borel–Cantelli.

LEMMA 4.8.

$$\limsup_{m\to\infty} |\mathcal{H}_m|^{1/m} \leqslant \frac{1}{z^*} \quad \textit{almost surely.}$$

Eventually, we obtain the desired upper box-counting dimension.

PROPOSITION 4.9.

$$\overline{\mathrm{BD}}(\Lambda) \leqslant \frac{\log z^*}{\log \alpha}.$$

Proof. Denote by $N(\alpha^m)$, the number of balls of radius of at most α^m needed to cover Λ. Then, for any $\varepsilon > 0$, $N(\alpha^m) \leqslant |\mathcal{H}_m| \leqslant (1/z^* + \varepsilon)^m$ almost surely for sufficiently large m. Therefore,

$$\overline{\mathrm{BD}}(\Lambda) = \limsup_{m\to\infty} -\frac{\log N(\alpha^m)}{\log \alpha^m} \leqslant \limsup_{m\to\infty} -\frac{\log(1/z^*+\varepsilon)^m}{\log \alpha^m} = -\frac{\log(1/z^*+\varepsilon)}{\log \alpha}.$$

Letting $\varepsilon \to 0$ proves the claim. □

4.2.2. *Lower bound for Hausdorff dimension.* In this section, we will show that $\log z^*/\log\alpha$ is also a lower bound for the Hausdorff dimension of Λ. From this, we may then conclude existence of the box-counting dimension since $\mathrm{HD}(\Lambda) \leqslant \underline{\mathrm{BD}}(\Lambda) \leqslant \overline{\mathrm{BD}}(\Lambda)$. The main idea of the proof follows [13]. This idea† is to construct a sequence of embedded Galton–Watson trees τ_r in the BRW such that the limit set Λ_{τ_r} of the Galton–Watson trees are subsets of the limit set Λ, see Section 6.3 in [13]. As r goes to infinity, we will have that $HD(\Lambda_{\tau_r}) \to HD(\lambda)$. This approximation property relies mainly on the facts that particles travel essentially along geodesics segments and that limit sets of multi-type Galton–Watson trees are well understood. Both facts hold still true for free products of finite groups and the proof of the lower bound is analogous to the one for free groups in [13], albeit technically more involved. The case of infinite factors need some extra care, since in this case particles do not necessarily travel along geodesics and infinite-type Galton–Watson processes are not so easy to handle. To bypass these difficulties, we approximate the infinite factors by increasing sequence of finite subgraphs. These subgraphs $\mathcal{X}_i^{(d)}$ are the subgraphs induced by the balls $B_i(d) := \{y \in \Gamma_i \mid l(y) \leqslant d\}$, $d \geqslant 1$. Letting $d \to \infty$ will give the optimal bound $\log z^*/\log\alpha$.

We add an additional vertex † to $\mathcal{X}_i^{(d)}$, the 'tomb', such that all edges in $\mathcal{X}_i$ exiting $B_i(d)$ now lead to the tomb. The random walk $(Y_n^{(i,d)})_{n\in\mathbb{N}_0}$ on $\mathcal{X}_i^{(d)}$ behaves like the random walk on Γ_i, with the exception that a particle leaving $B_i(d)$ dies. We now build the free product $\mathcal{X}^{(d)}$ from the $\mathcal{X}_i^{(d)}$, whose vertices are given by the set

$$\left\{ x_1 \dots x_n \in \Gamma \,\middle|\, n \in \mathbb{N}, x_j \in \bigcup_{i\in\mathcal{I}} \mathcal{X}_i^{(d)} \setminus \{e_i, \dagger\}, x_j \in \mathcal{X}_i^{(d)} \Rightarrow x_{j+1} \notin \mathcal{X}_i^{(d)} \right\} \cup \{e, \dagger\},$$

where † symbolizes the tomb. We identify $x \in \mathcal{X}^{(d)}$ with the corresponding element in Γ. Analogously to Section 2.1, we lift the random walks on the graphs $\mathcal{X}_i^{(d)}$ to a random walk $(X_n^{(d)})_{n\in\mathbb{N}_0}$ on $\mathcal{X}^{(d)}$ and define the associated BRW. We use the same notation (for Green functions, generating functions, etc.) as for the random walk on Γ itself but for reason of distinguishing we add superscripts '(d)', that is, we write, for example, $G^{(d)}(x,y|z)$ for the

†In this section, the parameter r is not identified with $|\mathcal{I}|$ but is used as a parameter of the Galton–Watson trees τ_r as in [13].

corresponding Green function of the random walk on $\mathcal{X}^{(d)}$. All involved generating functions on $\mathcal{X}^{(d)}$ have radii of convergence of at least R.

For any $x, y \in \Gamma$, we define $\overline{x:y}$ to be the set of vertices $w \in \Gamma$ such that there is a geodesic from x to y which passes through w. For $u \in \Gamma$, $d(u, \overline{x:y})$ is defined as the minimal distance with respect to the graph metric of u to any element of $\overline{x:y}$. In the case of the coloured BRW on $\mathcal{X}^{(d)}$, let $Z_\infty^{(d)}(y|x)$ be the overall number of blue particles arriving and freezing at $y \in \mathcal{X}^{(d)}$ under the assumption that the BRW is started with one blue particle at x. For $r \in \mathbb{N}$, we write $Z_{\infty,r}^{(d)}(y|x)$ for the overall number of particles counted in $Z_\infty^{(d)}(y|x)$ whose trail remain within distance r to a geodesic from x to y. In other words, in all sites u with $d(u, \overline{x:y}) > r$ every blue particle is coloured red. In the following, we set $x_0 := x_1^{-1}$, for any $x = x_1 \dots x_m \in \mathcal{X}^{(d)}$. The proofs of the following two lemmas are similar to the ones of Lemma 4 and Proposition 7 in [**13**] and are therefore omitted.†

Lemma 4.10.

$$\lim_{r\to\infty} \inf_{x=x_1\dots x_m \in \mathcal{X}^{(d)}} \left(\frac{\prod_{j=1}^m \mathbb{E} Z_{\infty,r}^{(d)}(x_1\dots x_j | x_1 \dots x_{j-1})}{\mathbb{E} Z_\infty^{(d)}(x|e)} \right)^{1/l(x)} = 1.$$

For $x \in \mathcal{X}^{(d)}$, we define the event $E^{(d)}(x)$ that among the particles counted in $Z_\infty^{(d)}(x|e)$ there is at least one particle whose trail has not entered $\Gamma_1^\times$ and enters the set

$$\{y \in \mathcal{X}^{(d)} \mid l(y) = l(x)\}$$

first at x. Obviously, $Z_\infty^{(d)}(x|e) \geqslant 1$ on the event $E^{(d)}(x)$ and hence $\mathbb{P}[E^{(d)}(x)] \leqslant \mathbb{E} Z_\infty^{(d)}(x|e)$.

Lemma 4.11.

$$\lim_{k\to\infty} \left(\min_{\substack{x=x_1\dots x_m \in \mathcal{X}^{(d)}: \\ m\in\mathbb{N}, x_1 \notin \Gamma_1, l(x)=k}} \frac{\mathbb{P}[E^{(d)}(x)]}{\mathbb{E} Z_\infty^{(d)}(x|e)} \right)^{1/k} = 1.$$

Analogously to (2.9) and (2.10), we define, for $i \in \mathcal{I}$ and $d \in \mathbb{N}$,

$$\begin{aligned} \mathcal{L}_i^{(d)+}(\lambda|z) &:= \sum_{x\in\Gamma_i^\times} L^{(d)}(e,x|\lambda)\, z^{l(x)} = \sum_{x\in\Gamma_i^\times} L_i^{(d)}(e_i, x|\xi_i^{(d)}(\lambda)) z^{l(x)}, \\ \mathcal{L}_i^{(d)}(\lambda|z) &:= \sum_{n\geqslant 1} \sum_{\substack{x=x_1\dots x_n \in \mathcal{X}^{(d)}: \\ \tau(x_1)=i}} L^{(d)}(e,x|\lambda)\, z^{l(x)} \\ &= \mathcal{L}_i^{(d)+}(\lambda|z) \left(1 + \sum_{j\in\mathcal{I}\setminus\{i\}} \mathcal{L}_j^{(d)}(\lambda|z) \right). \end{aligned} \tag{4.7}$$

Writing $\mathcal{L}^{(d)}(\lambda|z) := 1 + \sum_{i\in\mathcal{I}} \mathcal{L}_i^{(d)}(\lambda|z)$, we obtain analogously to Equation (4.1):

$$\mathcal{L}^{(d)}(\lambda|z) = \frac{1}{1 - \sum_{i\in\mathcal{I}} (\mathcal{L}_i^{(d)+}(\lambda|z)/(1+\mathcal{L}_i^{(d)+}(\lambda|z)))}.$$

†The reader may find all the details in the *arxiv.org* version (preprint) of this paper.

Since every function $\mathcal{L}_i^{(d)+}(\lambda|z)$ is convergent and strictly increasing, for all $z \geqslant 0$, there is some unique $z^*_{d,\mathcal{L}} > 0$ such that $\sum_{i\in\mathcal{I}} \mathcal{L}_i^{(d)+}(\lambda|z^*_{d,\mathcal{L}})/(1+\mathcal{L}_i^{(d)+}(\lambda|z^*_{d,\mathcal{L}})) = 1$. The radius of convergence of $\mathcal{L}^{(d)}(\lambda|z)$ is then given by $z^*_{d,\mathcal{L}}$.

We define for $k \in \mathbb{N}$

$$S_k^* := \{x_1 \dots x_s \in \mathcal{X}^{(d)} \mid s \in \mathbb{N}, l(x) = k, x_1 \notin \Gamma_1, x_s \in \Gamma_1\}.$$

Since we excluded the case $|\mathcal{I}| = 2 = |\Gamma_1| = |\Gamma_2|$, we have that $S_2^* \neq \emptyset$ and $S_3^* \neq \emptyset$. Therefore, $S_k^* \neq \emptyset$ for all $2 \leqslant k \in \mathbb{N}$.

Lemma 4.12.

$$\limsup_{k\to\infty} \left(\sum_{x\in S_k^*} \mathbb{P}[E^{(d)}(x)] \right)^{1/k} = \frac{1}{z^*_{d,\mathcal{L}}}.$$

Proof. By Lemma 4.11, we have $\mathbb{P}[E^{(d)}(x)] \geqslant (1-\varepsilon)^k \mathbb{E} Z_\infty^{(d)}(x|e)$ uniformly for all x with $l(x) = k$, if k is large enough. Recall also $\mathbb{P}[E^{(d)}(x)] \leqslant \mathbb{E} Z_\infty^{(d)}(x|e)$. Thus, it is sufficient to prove

$$\limsup_{k\to\infty} \left(\sum_{x\in S_k^*} \mathbb{E} Z_\infty^{(d)}(x|e) \right)^{1/k} = \frac{1}{z^*_{d,\mathcal{L}}}.$$

Since

$$\sum_{x\in S_k^*} \mathbb{E} Z_\infty^{(d)}(x|e) = \sum_{x\in S_k^*} F^{(d)}(e,x|\lambda) = \sum_{x\in S_k^*} \frac{G^{(d)}(e,e|\lambda)}{G^{(d)}(x,x|\lambda)} L^{(d)}(e,x|\lambda)$$

and $1 \leqslant G^{(d)}(x,x|\lambda) \leqslant G(x,x|\lambda) = G(e,e|\lambda) < \infty$, we have

$$\limsup_{k\to\infty} \left(\sum_{x\in S_k^*} L^{(d)}(e,x|\lambda) \right)^{1/k} = \limsup_{k\to\infty} \left(\sum_{x\in S_k^*} \mathbb{E} Z_\infty^{(d)}(x|e) \right)^{1/k}. \tag{4.8}$$

To determine the left-hand side of (4.8), we define further generating functions:

$$\mathcal{L}_{\neg 1,1}^{(d)}(\lambda|z) := \sum_{n\geqslant 2} \sum_{\substack{x=x_1\dots x_n \in \mathcal{X}^{(d)}: \\ x_1 \notin \Gamma_1^\times, x_n \in \Gamma_1^\times}} L^{(d)}(e,x|\lambda)\, z^{l(x)},$$

$$\mathcal{L}^{(d)*}(\lambda|z) := \sum_{n\geqslant 1} \sum_{\substack{x=x_1\dots x_n \in \mathcal{X}^{(d)}: \\ x_1,x_2,\dots,x_n \notin \Gamma_1}} L^{(d)}(e,x|\lambda)\, z^{l(x)}.$$

For $k \in \mathbb{N}$, the coefficient of z^k in $\mathcal{L}_{\neg 1,1}^{(d)}(\lambda|z)$ is just $\sum_{x\in S_k^*} L^{(d)}(e,x|\lambda)$. Due to Equation (4.7), we have

$$\mathcal{L}_1^{(d)}(\lambda|z) = \mathcal{L}_1^{(d)+}(\lambda|z) \cdot \left(1 + \sum_{i\in\mathcal{I}\setminus\{1\}} \mathcal{L}_i^{(d)}(\lambda|z)\right),$$

and hence the function $\mathcal{L}_{\neg 1}^{(d)}(\lambda|z) := 1 + \sum_{i\in\mathcal{I}\setminus\{1\}} \mathcal{L}_i^{(d)}(\lambda|z)$ must have the same radius of convergence as $\mathcal{L}^{(d)}(\lambda|z)$, which is $z^*_{d,\mathcal{L}}$. Moreover, we have the following relations:

$$\mathcal{L}_{\neg 1}^{(d)}(\lambda|z) = 1 + \mathcal{L}_{\neg 1,1}^{(d)}(\lambda|z)(1+\mathcal{L}^{(d)*}(\lambda|z)) + \mathcal{L}^{(d)*}(\lambda|z),$$
$$\mathcal{L}_{\neg 1,1}^{(d)}(\lambda|z) \geqslant \mathcal{L}^{(d)*}(\lambda|z) \cdot \mathcal{L}_1^{(d)+}(\lambda|z).$$

Since $\mathcal{L}^{(d)}_{\neg 1,1}(\lambda|z), \mathcal{L}^{(d)*}(\lambda|z) \leqslant \mathcal{L}^{(d)}_{\neg 1}(\lambda|z)$, the function $\mathcal{L}^{(d)}_{\neg 1,1}$ also has a radius of convergence of $z^*_{d,\mathcal{L}}$. □

Now we show that $z^*_{d,\mathcal{L}}$ tends to z^* as $d \to \infty$. Since $z^*_{d,\mathcal{L}}$ is strictly decreasing as d grows and due to

$$\lim_{d\to\infty} L^{(d)}(e,x|\lambda) = L(e,x|\lambda) = F(e,x|\lambda) \tag{4.9}$$

we have $z_\infty = \lim_{d\to\infty} z^*_{d,\mathcal{L}} \geqslant z^*$. Assume now, for a moment, that $z^* < z_\infty$. Then $\mathcal{F}^+_i(\lambda|z_\infty) < \infty$: indeed, assume that $\lim_{d\to\infty} \mathcal{L}^{(d)+}_j(\lambda|z_\infty) = \mathcal{F}^+_j(\lambda|z_\infty) = \infty$, for some $j \in \mathcal{I}$. Then we get the following contradiction:

$$1 = \lim_{d\to\infty} \sum_{i\in\mathcal{I}} \frac{\mathcal{L}^{(d)+}_i(\lambda|z^*_{d,\mathcal{L}})}{1+\mathcal{L}^{(d)+}_i(\lambda|z^*_{d,\mathcal{L}})} \geqslant \lim_{d\to\infty} \sum_{i\in\mathcal{I}} \frac{\mathcal{L}^{(d)+}_i(\lambda|z_\infty)}{1+\mathcal{L}^{(d)+}_i(\lambda|z_\infty)} > 1, \tag{4.10}$$

since $\mathcal{L}^{(d)+}_j(\lambda|z_\infty)/(1+\mathcal{L}^{(d)+}_j(\lambda|z_\infty))$ is arbitrarily close to 1, if d is large enough. Hence, $\mathcal{F}^+_i(\lambda|z_\infty) < \infty$. Now $z_\infty > z^*$ yields the following contradiction:

$$\begin{aligned} 1 &= \lim_{d\to\infty} \sum_{i\in\mathcal{I}} \frac{\mathcal{L}^{(d)+}_i(\lambda|z^*_{d,\mathcal{L}})}{1+\mathcal{L}^{(d)+}_i(\lambda|z^*_{d,\mathcal{L}})} \geqslant \limsup_{d\to\infty} \sum_{i\in\mathcal{I}} \frac{\mathcal{L}^{(d)+}_i(\lambda|z_\infty)}{1+\mathcal{L}^{(d)+}_i(\lambda|z_\infty)} \\ &= \sum_{i\in\mathcal{I}} \frac{\mathcal{F}^+_i(\lambda|z_\infty)}{1+\mathcal{F}^+_i(\lambda|z_\infty)} > 1. \end{aligned}$$

which produces a contradiction. Thus,

$$\lim_{d\to\infty} z^*_{d,\mathcal{L}} = z^*. \tag{4.11}$$

Let $2 \leqslant k \in \mathbb{N}$ arbitrary, but fixed. Similar to [**13**] we define an embedded Galton–Watson process of the BRW on the free product $\mathcal{X}^{(d)}$. For $n \in \mathbb{N}_0$, we define generations gen(n) S^*_{nk} and distinguished particles ζ_x associated to vertices $x \in$ gen(n) inductively as follows:

(1) gen$(0) := \{e\}$ consists of one particle ζ_e located at e.
(2) $y \in S^*_{(n+1)k}$ belongs to gen$(n+1)$ if and only if there exists a distinguished particle ζ_x in gen(n) such that some of its offspring particles counted in $Z^{(d)}_\infty(y|x)$ has a trail which
  (a) remains in the set
  $$\Gamma(x) := \{y \in \Gamma \mid y \text{ has the form } xw_1\dots w_s \text{ with } w_1 \notin \Gamma_1, s \geqslant 1\} \cup \{x\},$$
  (b) hits the set $\{w \in \mathcal{X}^{(d)} \mid l(w) = (n+1)k\}$ first at y.
(3) The first particle hitting $y \in S^*_{(n+1)k}$ becomes the distinguished particle ζ_y.

Let ϕ_n denote the number of particles in generation n. Since we have the same offspring distribution at every $x \in S^*_{nk}$, $(\phi_n)_{n\geqslant 0}$ defines a Galton–Watson process with mean $M_{d,k}$.

Corollary 4.13.

$$\limsup_{k\to\infty} M^{1/k}_{d,k} = \frac{1}{z^*_{d,\mathcal{L}}}.$$

Proof. The claim follows directly with Lemma 4.12 since $M_{d,k} = \sum_{x\in S^*_k} \mathbb{P}[E^{(d)}(x)]$. □

Applying Hawkes' Theorem as in Corollary 7 in [**13**] together with Equation (4.11) yields the following corollary.

Corollary 4.14. *We have* $\mathrm{HD}(\Lambda \cap \Omega_\infty) \geqslant \log z^*/\log \alpha$.

Proof of Theorem 3.5. The following chains of inequalities summarize the previous results and finish the proof of the theorem:

$$\frac{\log z^*}{\log \alpha} \leqslant \mathrm{HD}(\Lambda) \leqslant \underline{\mathrm{BD}}(\Lambda) \leqslant \overline{\mathrm{BD}}(\Lambda) \leqslant \frac{\log z^*}{\log \alpha},$$
$$\frac{\log z^*}{\log \alpha} \leqslant \mathrm{HD}(\Lambda \cap \Omega_\infty) \leqslant \underline{\mathrm{BD}}(\Lambda \cap \Omega_\infty) \leqslant \overline{\mathrm{BD}}(\Lambda \cap \Omega_\infty) \leqslant \overline{\mathrm{BD}}(\Lambda) \leqslant \frac{\log z^*}{\log \alpha}. \qquad \square$$

Proof of Corollary 3.7. It is well known that the Hausdorff dimension of a countable union $\bigcup_i B_i$ of sets $B_i \subseteq \Omega$ equals the supremum of the Hausdorff dimensions of the single sets B_i. Thus,

$$\mathrm{HD}(\Lambda \cap \Omega_i) = \sup_{x\in\Gamma:\tau(x)\neq i} \mathrm{HD}(\Lambda \cap x\Omega_i^{(0)}) \leqslant \sup_{x\in\Gamma:\tau(x)\neq i} \overline{\mathrm{BD}}(\Lambda \cap x\Omega_i^{(0)}).$$

For arbitrary, but fixed $x \in \Gamma$ with $\tau(x) \neq i$, denote by $\mathcal{H}_m^{(x)}$ the vertices $y \in x\Gamma_i$ with $l(y) = l(x) + m$, which are visited by the BRW. Therefore,

$$\mathbb{E}|\mathcal{H}_m^{(x)}| \leqslant \sum_{y\in\Gamma_i:l(y)=m} F(e, xy|\lambda) = F(e, x|\lambda) \sum_{y\in\Gamma_i:l(y)=m} F(e, y|\lambda).$$

Define

$$\mathcal{F}_{x,i}^{+}(\lambda|z) := F(e, x|\lambda) \sum_{m\geqslant 1} \sum_{y\in\Gamma_i:l(y)=m} F(e, y|\lambda)\, z^m.$$

The radius of convergence of $\mathcal{F}_{x,i}^{+}(\lambda|z)$ is obviously $R(\mathcal{F}_i^{+})$. Therefore, Lemma 4.7 yields $\limsup_{m\to\infty} (\mathbb{E}|\mathcal{H}_m^{(x)}|)^{1/m} \leqslant 1/R(\mathcal{F}_i^{+}) < 1/z^*$. The rest follows analogously to the proofs of Lemma 4.8 and Proposition 4.9. $\square$

4.3. *Proof of Theorem* 3.8 *and Corollary* 3.9

In order to prove Theorem 3.8, we can follow the argumentation of the proof of Theorem 3.5. For this purpose, we define, for $m \in \mathbb{N}$ and $i \in \mathcal{I}$,

$$S_i(m) := |\{x \in \Gamma_i \mid l(x) = m\}|, \quad S(m) := |\{x \in \Gamma \mid l(x) = m\}|,$$
$$S^{(i)}(m) := |\{x = x_1 \dots x_n \in S(m) \mid n \in \mathbb{N}, x_1 \in \Gamma_i\}|.$$

To cover Ω by balls of radius α^m we need at least $S(m-1)$ balls: indeed, for all $x, y \in \Gamma$, $x \neq y$, with $l(x) = l(y) = m - 1$, we can choose $v_x \in \Gamma_*^{\times} \setminus \Gamma_{\tau(x)}$ and $v_y \in \Gamma_*^{\times} \setminus \Gamma_{\tau(y)}$; then all balls of the form $B(\omega_1, \alpha^m)$ and $B(\omega_2, \alpha^m)$, where xv_x lies in the ω_1-component of $\mathcal{X} \setminus B_{m-1}$ and yv_y in the ω_2-component, do not intersect. Apparently, we need at most $S(m)$ balls of radius α^m to cover Ω. Obviously, the same holds for covering Ω_∞. We are now interested in the behaviour of $S(m)^{1/m}$ as $m \to \infty$. We define

$$\mathcal{S}_i^{+}(z) := \sum_{m\geqslant 1} S_i(m)\, z^m, \quad \mathcal{S}_i(z) := \sum_{m\geqslant 1} S^{(i)}(m)\, z^m,$$
$$\mathcal{S}(z) := \sum_{m\geqslant 0} S(m)\, z^m = 1 + \sum_{i\in\mathcal{I}} \mathcal{S}^{(i)}(z).$$

Analogously to the computations in Section 4.2.1 (we just replace the functions $\mathcal{F}_i^{+}(\lambda|z)$, $\mathcal{F}_i(\lambda|z)$ and $\mathcal{F}(\lambda|z)$ by the functions $\mathcal{S}_i^{+}(z)$, $\mathcal{S}_i(z)$ and $\mathcal{S}(z)$) we obtain

$$\mathcal{S}(z) = \frac{1}{1 - \sum_{i\in\mathcal{I}} (\mathcal{S}_i^{+}(z)/(1 + \mathcal{S}_i^{+}(z)))}. \tag{4.12}$$

Lemma 4.15.

$$\lim_{m\to\infty} S(m)^{1/m} = \frac{1}{z^*_{\mathcal{S}}} < 1,$$

where $z^*_{\mathcal{S}}$ is the smallest positive real number with

$$\sum_{i\in\mathcal{I}} \frac{\mathcal{S}_i^+(z^*_{\mathcal{S}})}{1+\mathcal{S}_i^+(z^*_{\mathcal{S}})} = 1.$$

Proof. Obviously, $R(\mathcal{S}) \leqslant R(\mathcal{F}) < 1$ since $F(e,x|\lambda) < 1$, for all $x \in \Gamma \setminus \{e\}$. The equation $R(\mathcal{S}) = z^*_{\mathcal{S}}$ follows now analogously to the proof of Lemma 4.7. This yields

$$\limsup_{m\to\infty} S(m)^{1/m} = \frac{1}{z^*_{\mathcal{S}}} = \frac{1}{R(\mathcal{S})} > 1.$$

Thus, it is sufficient to prove convergence of $S(m)^{1/m}$ as $m \to \infty$. By transitivity of Γ, we have $S(m)S(n) \geqslant S(m+n)$, for all $m,n \in \mathbb{N}$. Therefore, $\log S(m) + \log S(n) \geqslant \log S(m+n)$, that is, $(\log S(m))_{m\in\mathbb{N}}$ forms a subadditive sequence. By Fekete's Lemma, $(1/m)\log S(m) = \log S(m)^{1/m}$ converges to some constant s, that is, $S(m)^{1/m}$ converges to e^s, which must equal $1/z^*_{\mathcal{S}}$. □

Remark 4.16. One can show analogously to Lemma 4.7 that $z^*_{\mathcal{S}} < R(\mathcal{S}_i^+)$, where $R(\mathcal{S}_i^+)$ is the radius of convergence of $\mathcal{S}_i^+(z)$. In particular, $z^*_{\mathcal{S}}$ is the radius of convergence of $\mathcal{S}(z)$.

We can conclude by giving a formula for $\mathrm{BD}(\Omega)$ and observing that the box-counting dimension of Ω results from the dimension of Ω_∞.

Proposition 4.17.

$$\mathrm{BD}(\Omega) = \mathrm{BD}(\Omega_\infty) = \frac{\log z^*_{\mathcal{S}}}{\log \alpha}.$$

Proof. Recall the remarks at the beginning of this section concerning the minimal and maximal number of balls needed to cover Ω_∞. This yields

$$\begin{aligned}\underline{\mathrm{BD}}(\Omega) \geqslant \underline{\mathrm{BD}}(\Omega_\infty) &\geqslant \liminf_{m\to\infty} -\frac{\log S(m-1)}{\log \alpha^m} \\ &= \liminf_{m\to\infty} -\frac{\log S(m-1)^{1/(m-1)}}{\log \alpha}\frac{m-1}{m} = \frac{\log z^*_{\mathcal{S}}}{\log \alpha}.\end{aligned}$$

Analogously,

$$\overline{\mathrm{BD}}(\Omega_\infty) \leqslant \overline{\mathrm{BD}}(\Omega) \leqslant \limsup_{m\to\infty} -\frac{\log S(m)}{\log \alpha^m} = \limsup_{m\to\infty} -\frac{\log S(m)^{1/m}}{\log \alpha} = \frac{\log z^*_{\mathcal{S}}}{\log \alpha}.$$

Both inequality chains together yield the formula for the box-counting dimension. □

Finally, we can prove the formula for the Hausdorff dimensions of Ω and Ω_∞.

Proof of Theorem 3.8. It is sufficient to show that $\mathrm{HD}(\Omega_\infty) \geqslant \log z^*_{\mathcal{S}}/\log\alpha$. Define for $d,k \in \mathbb{N}$ and $i \in \mathcal{I}$

$$\begin{aligned}
S_i^{(d)+}(k) &= |\{x \in \mathcal{X}_i^{(d)} \setminus \{\dagger\} \mid l(x)=k\}|,\ S^{(d)}(k) = |\{x \in \mathcal{X}^{(d)} \setminus \{\dagger\} \mid l(x)=k\}|,\\
S_i^{(d)}(k) &= |\{x_1\dots x_s \in \mathcal{X}^{(d)} \setminus \{\dagger\} \mid s \in \mathbb{N}, l(x)=k, x_1 \in \Gamma_i\}|,\\
S_{\neg 1}^{(d)}(k) &= |\{x_1\dots x_s \in \mathcal{X}^{(d)} \setminus \{\dagger\} \mid s \in \mathbb{N}, l(x)=k, x_1 \notin \Gamma_1\}|,\\
S_{\neg 1,\neg 1}^{(d)}(k) &= |\{x_1\dots x_s \in \mathcal{X}^{(d)} \setminus \{\dagger\} \mid s \in \mathbb{N}, l(x)=k, x_1, x_s \notin \Gamma_1\}|,\\
S_{\neg 1,1}^{(d)}(k) &= |\{x_1\dots x_s \in \mathcal{X}^{(d)} \setminus \{\dagger\} \mid s \in \mathbb{N}, l(x)=k, x_1 \notin \Gamma_1, x_s \in \Gamma_1\}|.
\end{aligned}$$

The associated generating functions are given by

$$\begin{aligned}
\mathcal{S}_i^{(d)+}(z) &= \sum_{k\geqslant 1} S_i^{(d)+}(k)\, z^k, \quad \mathcal{S}^{(d)}(z) = \sum_{k\geqslant 0} S^{(d)}(k)\, z^k,\\
\mathcal{S}_i^{(d)}(z) &= \sum_{k\geqslant 1} S_i^{(d)}(k)\, z^k, \quad \mathcal{S}_{\neg 1}^{(d)}(z) = \sum_{k\geqslant 1} S_{\neg 1}^{(d)}(k)\, z^k,\\
\mathcal{S}_{\neg 1,\neg 1}^{(d)}(z) &= \sum_{k\geqslant 1} S_{\neg 1,\neg 1}^{(d)}(k)\, z^k, \quad \mathcal{S}_{\neg 1,1}^{(d)}(z) = \sum_{k\geqslant 1} S_{\neg 1,1}^{(d)}(k)\, z^k.
\end{aligned}$$

Once again, we write

$$\mathcal{S}^{(d)}(z) = \frac{1}{1-\sum_{i\in\mathcal{I}}(\mathcal{S}_i^{(d)+}(z)/(1+\mathcal{S}_i^{(d)+}(z)))}$$

and obtain

$$\mathcal{S}_2^{(d)+}(z)\mathcal{S}_1^{(d)+}(z)\mathcal{S}_{\neg 1,\neg 1}^{(d)}(z)\mathcal{S}_1^{(d)+}(z) \leqslant \mathcal{S}_{\neg 1,1}^{(d)}(z) = \mathcal{S}_{\neg 1}^{(d)}(z) - \mathcal{S}_{\neg 1,\neg 1}^{(d)}(z).$$

Thus, $\mathcal{S}_{\neg 1,1}^{(d)}(z)$ and $\mathcal{S}_{\neg 1}^{(d)}(z)$ have the same radius of convergence. Moreover,

$$\mathcal{S}_2^{(d)+}(z)\mathcal{S}_1^{(d)}(z) \leqslant \mathcal{S}_{\neg 1}^{(d)}(z) = \mathcal{S}^{(d)}(z) - \mathcal{S}_1^{(d)}(z) - 1.$$

That is, $\mathcal{S}_{\neg 1,1}^{(d)}(z)$ and $\mathcal{S}^{(d)}(z)$ have the same radius of convergence, which is given by $z^*_{d,\mathcal{S}}$, the smallest positive solution satisfying

$$1 = \sum_{i\in\mathcal{I}} \frac{\mathcal{S}_i^{(d)+}(z)}{1+\mathcal{S}_i^{(d)+}(z)}.$$

Since $z^*_{d,\mathcal{S}}$ is strictly decreasing as $d\to\infty$, we have that $\lim_{d\to\infty} z^*_{d,\mathcal{S}} = z^*_{\mathcal{S}}$. This can be seen by contradiction. Indeed, if $\lim_{d\to\infty} z^*_{d,\mathcal{S}} = z^*_{\infty,\mathcal{S}} > z^*_{\mathcal{S}}$, then $\mathcal{S}_i^+(z^*_{\infty,\mathcal{S}}) < \infty$, for all $i \in \mathcal{I}$ (this is proved analogously as explained in Equation (4.10)) and therefore

$$1 = \lim_{d\to\infty} \sum_{i\in\mathcal{I}} \frac{\mathcal{S}_i^{(d)+}(z^*_{d,\mathcal{S}})}{1+\mathcal{S}_i^{(d)+}(z^*_{d,\mathcal{S}})} \geqslant \lim_{d\to\infty} \sum_{i\in\mathcal{I}} \frac{\mathcal{S}_i^{(d)+}(z^*_{\infty,\mathcal{S}})}{1+\mathcal{S}_i^{(d)+}(z^*_{\infty,\mathcal{S}})} = \sum_{i\in\mathcal{I}} \frac{\mathcal{S}_i^{+}(z^*_{\infty,\mathcal{S}})}{1+\mathcal{S}_i^{+}(z^*_{\infty,\mathcal{S}})} > 1,$$

a contradiction. Thus,

$$(S_{\neg 1,1}^{(d)}(k))^{1/k} \xrightarrow{k\to\infty} \frac{1}{z^*_{d,\mathcal{S}}} \xrightarrow{d\to\infty} \frac{1}{z^*_{\mathcal{S}}}.$$

We can embed a 'deterministic' Galton–Watson tree into the free product analogously to Section 4.2.2, where each generation has exactly $S_{\neg 1,1}^{(d)}(k)$ descendants. By Hawkes's Theorem, the Hausdorff dimension of the boundary of the embedded tree is bounded from below by $\log z^*_{d,\mathcal{S}}/\log\alpha$, and therefore $\mathrm{HD}(\Omega_\infty) \geqslant \log z^*_{\mathcal{S}}/\log\alpha$. □

Proof of Corollary 3.9. Analogously to the proof of Corollary 3.7 and by Remark 4.16, we can use the property $\mathrm{HD}(\cup_i B_i) = \sup_i \mathrm{HD}(B_i)$ for all countable unions of sets $B_i \subseteq \Omega$ in order to show that

$$\mathrm{HD}(\Omega_i) = \sup_{x\in\Gamma:\tau(x)\neq i} \mathrm{HD}(x\Omega_i^{(0)}) \leqslant \overline{\mathrm{BD}}(\Omega_i^{(0)}) < \mathrm{BD}(\Omega_\infty) = \mathrm{HD}(\Omega_\infty). \qquad \square$$

4.4. *Proof of Theorem* 3.10

Proof of Theorem 3.10(1). In the following, we write $z^* = z^*(\lambda)$ in order to distinguish the solutions of (3.1) for different values of λ. Note that $z^*(\lambda_1) > z^*(\lambda_2)$, if $\lambda_1 < \lambda_2$. This implies the strictly increasing behaviour of Φ in the interval $(1, R]$. Recall that the BRW does almost surely *not* survive in the limit case $\lambda = 1$, yielding $\Phi(1) = 0$. Moreover, if $\lambda > R$, then the BRW is recurrent and thus $\mathrm{HD}(\Lambda) = \mathrm{HD}(\Omega)$. $\square$

The proof of Theorem 3.10(2) splits up into the following two lemmas:

Lemma 4.18. *The function Φ is continuous in $[1,\infty) \setminus \{R\}$ and continuous from the left at $\lambda = R$.*

Proof. In order to prove continuity of Φ, it is sufficient to prove continuity of the mapping $\lambda \mapsto z^* = z^*(\lambda)$. First, we prove continuity from the left at $\lambda_0 \in (1,\infty)$. For this purpose, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence of strictly increasing real numbers with $\lambda_n < \lambda_0$ and $\lim_{n\to\infty} \lambda_n = \lambda_0$. We use a proof by contradiction. Assume $z_0 := \lim_{n\to\infty} z^*(\lambda_n) > z^*(\lambda_0)$ (by simple domination arguments, $z^*(\lambda_n)$ cannot be less than $z^*(\lambda_0)$). We have that $z^*(\lambda_n)$ is strictly decreasing and

$$\mathcal{F}_i^+(\lambda_n|z^*(\lambda_0)) + \xi_i(1)(z_0 - z^*(\lambda_0)) \leqslant \mathcal{F}_i^+(\lambda_n|z_0) < \infty.$$

Here we used the fact that the coefficient of z in $\mathcal{F}_i^+(\lambda|z)$ is at least $\xi_i(1)$. We set $c := \xi_i(1)(z_0 - z^*(\lambda_0))$. Since $f(x)/(1+f(x))$ is strictly increasing in $[1,\infty)$ if $f(x)$ is a strictly increasing function on $[1,\infty)$ we obtain the following contradiction:

$$\begin{aligned}
1 &= \lim_{n\to\infty} \sum_{i\in\mathcal{I}} \frac{\mathcal{F}_i^+(\lambda_n|z^*(\lambda_n))}{1+\mathcal{F}_i^+(\lambda_n|z^*(\lambda_n))} \\
&\geqslant \limsup_{n\to\infty} \sum_{i\in\mathcal{I}} \frac{\mathcal{F}_i^+(\lambda_n|z_0)}{1+\mathcal{F}_i^+(\lambda_n|z_0)} \\
&\geqslant \limsup_{n\to\infty} \sum_{i\in\mathcal{I}} \frac{\mathcal{F}_i^+(\lambda_n|z^*(\lambda_0)) + c}{1+\mathcal{F}_i^+(\lambda_n|z^*(\lambda_0)) + c} \\
&= \sum_{i\in\mathcal{I}} \frac{\mathcal{F}_i^+(\lambda_0|z^*(\lambda_0)) + c}{1+\mathcal{F}_i^+(\lambda_0|z^*(\lambda_0)) + c} \\
&> \sum_{i\in\mathcal{I}} \frac{\mathcal{F}_i^+(\lambda_0|z^*(\lambda_0))}{1+\mathcal{F}_i^+(\lambda_0|z^*(\lambda_0))} = 1.
\end{aligned}$$

Thus, $\lim_{n\to\infty} z^*(\lambda_n) = z^*(\lambda_0)$.

Since $\mathrm{HD}(\Lambda) = \mathrm{HD}(\Omega)$ for all $\lambda > R$, it remains to prove continuity from the right for $\lambda_0 \in (1, R)$. We make a case distinction whether $\xi_i(\lambda_0) < 1$ or not. If $\xi_i(\lambda_0) < 1$, then $\mathcal{F}_i^+(\lambda_0 + \delta|1) < \infty$, for all $\delta > 0$ with $\xi_i(\lambda_0 + \delta) < 1$ according to (4.4). Moreover, $z^*(\lambda_0) < 1$. Therefore, continuity from the right follows directly from the Implicit Function Theorem, since $z^* = z^*(\lambda)$ is given by the equation

$$1 = \sum_{i\in\mathcal{I}} \frac{\mathcal{F}_i^+(\lambda \mid z^*(\lambda))}{1+\mathcal{F}_i^+(\lambda \mid z^*(\lambda))}.$$

We note that the derivative $\partial\mathcal{F}_i^+(\lambda|z)/\partial z$ evaluated at $z = z^*(\lambda)$ is positive and finite, since $z^*(\lambda)$ is strictly smaller than the radius of convergence of $\mathcal{F}_i^+(\lambda|z)$; see Lemma 4.7.

Now we turn to the case $\xi_i(\lambda_0) \geqslant 1$. Let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence of strictly decreasing real numbers with $\lambda_0 < \lambda_n < R$ and $\lim_{n\to\infty} \lambda_n = \lambda_0$. Assume $z_0 := \lim_{n\to\infty} z^*(\lambda_n) < z^*(\lambda_0)$ (by simple domination arguments, $z^*(\lambda_n)$ cannot be larger than $z^*(\lambda_0)$). Observe that $z^*(\lambda_n)$ is strictly increasing. By (4.6), there is $C := \sqrt{1+\xi_2(1)/(2|\mathrm{supp}(\mu_1)|)} > 1$ such that $Cz^*(\lambda_n) \leqslant R(\mathcal{F}_i^+)$, for all $n \in \mathbb{N}$. Choose $\tilde{C} \in (1, C)$ such that $\tilde{C}z_0 < z^*(\lambda_0)$ and choose $N \in \mathbb{N}$ large enough such that $\tilde{C}z^*(\lambda_n) \geqslant z_0$, for all $n \geqslant N$. Therefore,

$$\begin{aligned} 1 &= \lim_{n\to\infty} \sum_{i\in\mathcal{I}} \frac{\mathcal{F}_i^+(\lambda_n|z^*(\lambda_n))}{1+\mathcal{F}_i^+(\lambda_n|z^*(\lambda_n))} \\ &\leqslant \lim_{n\to\infty} \sum_{i\in\mathcal{I}} \frac{\mathcal{F}_i^+(\lambda_n|\tilde{C}z_0)}{1+\mathcal{F}_i^+(\lambda_n|\tilde{C}z_0)} = \sum_{i\in\mathcal{I}} \frac{\mathcal{F}_i^+(\lambda_0|\tilde{C}z_0)}{1+\mathcal{F}_i^+(\lambda_0|\tilde{C}z_0)} < 1, \end{aligned}$$

a contradiction. Consequently, $\lim_{n\to\infty} z^*(\lambda_n) = z^*(\lambda_0)$.

It remains to prove continuity from the right at $\lambda_0 = 1$. In this case, $\xi_i(1) < 1$. Once again $\mathcal{F}_i^+(\lambda_0+\delta|1) < \infty$ for all $\delta > 0$ with $\xi_i(\lambda_0+\delta) < 1$ according to (4.4). Let $(\lambda_n)_{n\in\mathbb{N}}$ be a strictly decreasing sequence of real numbers with limit 1. We write $z_0 = \lim_{n\to\infty} z^*(\lambda_n) \leqslant 1$. Then, for n large enough,

$$\begin{aligned} 1 &= \lim_{n\to\infty} \sum_{i\in\mathcal{I}} \frac{\mathcal{F}_i^+(\lambda_n|z^*(\lambda_n))}{1+\mathcal{F}_i^+(\lambda_n|z^*(\lambda_n))} \\ &\leqslant \lim_{n\to\infty} \sum_{i\in\mathcal{I}} \frac{\mathcal{F}_i^+(\lambda_n|z_0)}{1+\mathcal{F}_i^+(\lambda_n|z_0)} = \sum_{i\in\mathcal{I}} \frac{\mathcal{F}_i^+(1|z_0)}{1+\mathcal{F}_i^+(1|z_0)}. \end{aligned}$$

In order to finish the proof, we verify that $z^*(1) = 1$, from which $z_0 = z^*(1) = 1$ follows. Indeed, by Equation (4.4), we obtain

$$\sum_{i\in\mathcal{I}} \frac{\mathcal{F}_i^+(1|1)}{1+\mathcal{F}_i^+(1|1)} = \sum_{i\in\mathcal{I}} (1 - G_i(e_i, e_i|\xi_1(1))(1-\xi_i(1))).$$

From [**9**, Lemma 5.1], it follows that $1 - G_i(e_i, e_i|\xi_1(1))(1-\xi_i(1))$ is just the probability that a single random walk on Γ tends to an infinite word of the form $x_1x_2\ldots \in \Omega_\infty$ with $x_1 \in \Gamma_i^\times$, that is, the above sum equals 1. □

The next result completes the proof of Theorem 3.10(2):

LEMMA 4.19. *For all* $\lambda \in [1, R]$, $\mathrm{HD}(\Lambda) \leqslant \frac{1}{2}\mathrm{HD}(\Omega)$.

Proof. Define the function

$$\mathcal{F}^{(2)}(\lambda|z) := \sum_{x\in\Gamma} F(e, x|\lambda)^2\, z^{l(x)},$$

whose radius of convergence is denoted by z_2^*. The Cauchy–Schwarz Inequality then gives

$$\begin{aligned}\frac{1}{z^*} &= \limsup_{m\to\infty} \left(\sum_{x\in\Gamma: l(x)=m} F(e,x|\lambda)\right)^{1/m}\\ &\leqslant \limsup_{m\to\infty} \sqrt{\left(\sum_{x\in\Gamma: l(x)=m} F(e,x|\lambda)^2\right)^{1/m}} \cdot \limsup_{m\to\infty} \sqrt{\left(\sum_{x\in\Gamma: l(x)=m} 1^2\right)^{1/m}}\\ &= \sqrt{\frac{1}{z_2^*}}\cdot\sqrt{\frac{1}{z_{\mathcal{S}}^*}}.\end{aligned}$$

To prove the claim of the lemma it suffices (by the formulas given in Theorems 3.5 and 3.8) to show that $z_2^* \geqslant 1$. First,

$$\begin{aligned}\mathcal{F}^{(2)}(\lambda|1) &= \sum_{x\in\Gamma} F(e,x|\lambda)^2 = \frac{1}{G(e,e|\lambda)^2}\sum_{x\in\Gamma} G(e,x|\lambda)^2\\ &= \frac{1}{G(e,e|\lambda)^2}\sum_{x\in\Gamma}\left(\sum_{n\geqslant 0} p^{(n)}(e,x)\lambda^n\right)^2.\end{aligned}$$

For given $x\in\Gamma$, the coefficient of λ^n in the inner squared sum can (by symmetry) be rewritten as

$$\frac{1}{G(e,e|\lambda)^2}\sum_{m=0}^{n} p^{(m)}(e,x)p^{(n-m)}(x,e). \tag{4.13}$$

Thus, every path $[x_0=e,x_1,\dots,x_n=e]$ of length n (consisting of $n+1$ vertices) from e to e is counted $n+1$ times, since every x_i can play the role of x in Equation (4.13). That is,

$$\mathcal{F}^{(2)}(\lambda|z) = \frac{1}{G(e,e|\lambda)^2}\sum_{n\geqslant 0} p^{(n)}(e,e)\cdot(n+1)\cdot\lambda^n = \frac{\lambda\, G'(e,e|\lambda)}{G(e,e|\lambda)^2} + \frac{1}{G(e,e|\lambda)}.$$

From this follows $z_2^*\geqslant 1$, whenever $\lambda<R$ or $G'(e,e|R)<\infty$, and thus $\mathrm{HD}(\Lambda)\leqslant \frac{1}{2}\mathrm{HD}(\Omega)$, for $\lambda<R$. By Lemma 4.18, the proposed inequality holds (due to continuity from the left) also in the case $\lambda=R$. □

In order to prove Theorem 3.10(3), we start with the following lemma:

Lemma 4.20. *For all $i\in\mathcal{I}$, $G_i'(e_i,e_i|\xi_i(R))<\infty$.*

Proof. From [**25**, Proposition 9.18], it follows $\xi_i(R)\leqslant R_i$, where R_i is the radius of convergence of $G_i(e_i,e_i|z)$. If $\xi_i(R)<R_i$, then the claim of the lemma is obvious. Assume now that $\xi_i(R)=R_i$. Then, by Woess [**25**, Lemma 17.1.(a)], $R\,G(e,e|R)=R_i\,G_i(e_i,e_i|R_i)/\alpha_i$. Therefore, $G_i(e_i,e_i|R_i)<\infty$ since $G(e,e|R)<\infty$ by non-amenability of Γ. If $G_i'(e_i,e_i|R_i)=\infty$ would hold, we would get a contradiction to $\xi_i(R)=R_i$ by Woess [**25**, Equation (9.14), Theorem 9.22, Lemma 17.1(a)]. □

Let us remark that $F_i'(e_i,x|\xi_i(R)) = F_i'(x,e_i|\xi_i(R)) < \infty$, for all $x \in \Gamma_*^\times$; this can be easily verified with the help of the inequality

$$\mu_i^{(n+|x|)}(e_i) \geqslant \mu_i^{(|x|)}(x) \cdot \mathbb{P}[Y_n^{(i)} = e_i, \forall m < n : Y_m^{(i)} \neq e_i \mid Y_0^{(i)} = x] \quad \text{for all } n \in \mathbb{N},$$

where $(Y_n^{(i)})_{n\in\mathbb{N}}$ is a random walk on Γ_i governed by μ_i. We proceed now with expanding the Green function $G(z) := G(e,e|z)$ in a neighbourhood of $z = R$. By [**25**, Propositon 17.4] and [**4**, Sections 3 & 4], we have

$$G(z) = \begin{cases} G(R) + g_1 \cdot \sqrt{R-z} + \mathbf{o}(\sqrt{R-z}) & \text{if } G'(R) = \infty, \\ G(R) - G'(R)\cdot(R-z) + \mathbf{o}(R-z) & \text{if } G'(R) < \infty. \end{cases}$$

We write in the following $c := \frac{1}{2}$, if $G'(R) = \infty$, and $c := 1$, otherwise. The next aim is to show that the functions $F(e,x|z)$, $x \in \Gamma \setminus \{e\}$, have the same expansions.

LEMMA 4.21. *For all $x \in \Gamma \setminus \{e\}$, there are constants $f_x \neq 0$ such that*

$$F(e,x|z) = F(e,x|R) + f_x \cdot (R-z)^c + \mathbf{o}((R-z)^c).$$

Proof. We consider the case $c = 1$ first. By Candellero and Gilch [**4**, Lemma 3.2], we have $0 < \xi_i'(R) < \infty$, that is, we can write

$$\xi_i(z) = \xi_i(R) - \xi_i'(R)\cdot(R-z) + \mathbf{o}(R-z).$$

In the following, we write $F_i(e_i,x|z) = \sum_{n\geqslant 1} f_n(x) z^n$, for $x \in \Gamma_i^\times$. Therefore,

$$F(e,x|z) = F_i(e_i,x|\xi_i(z)) = \sum_{n\geqslant 1} f_n(x)(\xi_i(R) - \xi_i'(R)\cdot(R-z) + \mathbf{o}(R-z))^n. \tag{4.14}$$

The coeffcient of $(R-z)$ is given by

$$-\xi_i'(R)\cdot\sum_{n\geqslant 1} n\cdot f_n(x)\cdot\xi_i(R)^{n-1} = -\xi_i'(R)F_i'(e_i,x|\xi_i(R)) \in (-\infty,0).$$

Recall that, for $x = x_1 \dots x_n \in \Gamma \setminus \{e\}$,

$$F(e,x_1\dots x_n|z) = \prod_{j=1}^n F_{\tau(x_j)}(e_{\tau(x_j)}, x_j \mid \xi_{\tau(x_j)}(z)).$$

Now, plugging the expansion (4.14) into the above formula gives us the coefficient of $(R-z)$:

$$\begin{aligned} f_x = \sum_{j=1}^n &-\xi_{\tau(x_j)}'(R) F_{\tau(x_j)}'(e_{\tau(x_j)}, x_j \mid \xi_{\tau(x_j)}(R)) \\ &\times \prod_{\substack{k=1,\\ k\neq j}}^n F_{\tau(x_k)}(e_{\tau(x_k)}, x_k \mid \xi_{\tau(x_k)}(R)) \in (-\infty,0). \end{aligned}$$

This yields the claim in the case $c = 1$.

We now turn to the case $c = \frac{1}{2}$. By Woess [**25**, Equation (9.20)], we have

$$\alpha_i z G(z) = \xi_i(z)\, G_i(\xi_i(z)). \tag{4.15}$$

Write $\xi_i(z) = \xi_i(R) + X_i(z)$ with $X_i(R) = 0$. Our aim is to show that $X_i(z)$ is of order $\sqrt{R-z}$, from which we can derive the proposed expansion of $F(e, x|z)$. We rewrite (4.15) as

$$\begin{aligned}\alpha_i(R-(R-z)) \cdot \left(G(R) + g_1\sqrt{R-z} + \mathbf{o}(\sqrt{R-z})\right) \\ = (\xi_i(R) + X_i(z)) \cdot \sum_{n\geqslant 0} \mu_i^{(n)}(e_i)(\xi_i(R) + X_i(z))^n.\end{aligned}$$

The constant term on the left-hand side of the equation is $\alpha_i\, R\, G(R)$, which equals the constant term on the right-hand side $\xi_i(R)G_i(\xi_i(R))$ by (4.15). The coefficient of $\sqrt{R-z}$ on the left-hand side is $\alpha_i R g_1 \neq 0$. The coefficient of $X_1(z)$ on the right-hand side is given by

$$\xi_i(R)G_i'(e_i, e_i|\xi_i(R)) + G_i(e_i, e_i|\xi_i(R)) \in (0, \infty).$$

Thus, $X_1(z) \sim \sqrt{R-z}$ as $z \uparrow R$, and therefore

$$F_i(e_i, x|\xi_i(z)) = \sum_{n\geqslant 1} f_n(x)\left(\xi_i(R) + \hat{\xi}_i \cdot \sqrt{R-z} + \mathbf{o}(R-z)\right)^n,$$

for some $\hat{\xi}_i < 0$. The rest follows analogously to the case $c = 1$ by replacing $(R-z)$ with $\sqrt{R-z}$. □

Consider now the following difference for $i \in \mathcal{I}$:

$$\begin{aligned}\mathcal{F}_i(R|z^*(R)) - \mathcal{F}_i(\lambda|z^*(\lambda)) &= \sum_{m\geqslant 1} z^*(R)^m \sum_{\substack{x\in\Gamma_i:\\ l(x)=m}} F(e, x|R) - \sum_{m\geqslant 1}(z^*(R) - (z^*(R) - z^*(\lambda)))^m \\ &\quad \times \sum_{\substack{x\in\Gamma_i:\\ |x|=m}} [F(e, x|R) + f_x(R-\lambda)^c + \mathbf{o}((R-\lambda)^c)] \\ &= \sum_{m\geqslant 1} z^*(R)^m \sum_{\substack{x\in\Gamma_i:\\ l(x)=m}} (-f_x(R-\lambda)^c - \mathbf{o}((R-\lambda)^c)) \\ &\quad + (z^*(R) - z^*(\lambda))\frac{\partial}{\partial z}\mathcal{F}_i^+(\lambda|z^*(R)) + \mathbf{o}(z^*(R) - z^*(\lambda)).\end{aligned}$$

Moreover,

$$\begin{aligned}0 &= \sum_{i\in\mathcal{I}} \frac{\mathcal{F}_i(R|z^*(R))}{1+\mathcal{F}_i(R|z^*(R))} - \sum_{i\in\mathcal{I}} \frac{\mathcal{F}_i(\lambda|z^*(\lambda))}{1+\mathcal{F}_i(\lambda|z^*(\lambda))} \\ &= \sum_{i\in\mathcal{I}}\sum_{n\geqslant 0} (-\mathcal{F}_i(\lambda|z^*(\lambda)))^{n+1} - (-\mathcal{F}_i(R|z^*(R)))^{n+1}.\end{aligned} \tag{4.16}$$

Write

$$(-\mathcal{F}_i(\lambda|z^*(\lambda)))^{n+1} - (-\mathcal{F}_i(R|z^*(R)))^{n+1} = (\mathcal{F}_i(R|z^*(R)) - \mathcal{F}_i(\lambda|z^*(\lambda))) \cdot g_n(\lambda), \tag{4.17}$$

where $g_n(R) \neq 0$, for every $n \in \mathbb{N}$. Plugging the decomposition of $\mathcal{F}_i(R|z^*(R)) - \mathcal{F}_i(\lambda|z^*(\lambda))$ into (4.16) and comparing all error terms yields in view of (4.17) the following behaviour:

$$z^*(R) - z^*(\lambda) \sim \begin{cases}\hat{C}_1 \cdot (R-\lambda) & \text{if } G'(R) < \infty, \\ \hat{C}_2 \cdot \sqrt{R-\lambda} & \text{if } G'(R) = \infty\end{cases}$$

for suitable constants $\hat{C}_1$ and $\hat{C}_2$, respectively. The statement (3) of Theorem 3.10 follows now from

$$\log z^*(\lambda) - \log z^*(R) = \log\left(1 - \frac{1}{z^*(R)}(z^*(R) - z^*(\lambda))\right)$$

and by the Taylor expansion of $\log(1-x)$ at $x = 0$.

4.5. *Proof of Corollary* 3.16

In a first step, we show the following lemma:

LEMMA 4.22.

$$\overline{\mathrm{BD}^{\mathrm{fin}}}(\Lambda) \leqslant -\frac{\log\theta}{\log\alpha} \quad \textit{and} \quad \mathrm{BD}^{\mathrm{fin}}(\Omega) = -\frac{\log\varrho}{\log\alpha}.$$

Proof. First, we define the matrices $M_0 = (m_0(i,j))_{i,j\in\mathcal{I}}$ and $D_0 = (d_0(i,j))_{i,j\in\mathcal{I}}$ by

$$m_0(i,j) := \begin{cases} \mathcal{F}_i^+(\lambda) & \text{if } i=j,\\ 0 & \text{otherwise,}\end{cases} \quad \text{and} \quad d_0(i,j) := \begin{cases} |\Gamma_i|-1 & \text{if } i=j,\\ 0 & \text{otherwise.}\end{cases}$$

For $m\in\mathbb{N}$, denote by $\mathcal{H}_m^{\mathrm{fin}}$ the random number of visited words of the form $w_1\dots w_m\in\Gamma$. Then

$$\mathbb{E}|\mathcal{H}_m^{\mathrm{fin}}| \leqslant \sum_{x\in\mathcal{H}_m^{\mathrm{fin}}} \mathbb{E}Z_\infty(x) \leqslant \sum_{x\in\Gamma:\|x\|=m} F(e,x|\lambda) = \mathbb{1}^T M_0 M^{m-1}\mathbb{1}$$

and

$$\hat{S}(m) = |\{x\in\Gamma \mid \|x\|=m\}| = \mathbb{1}^T D_0 D^{m-1}\mathbb{1}.$$

Let $u\in\mathbb{R}^r$ be an eigenvector with respect to the eigenvalue θ such that $u\geqslant\mathbb{1}$. Then:

$$\mathbb{E}|\mathcal{H}_m^{\mathrm{fin}}| \leqslant \begin{pmatrix}\mathcal{F}_1(\lambda)\\ \vdots \\ \mathcal{F}_r(\lambda)\end{pmatrix}^T M_1^{m-1}u \leqslant \begin{pmatrix}\mathcal{F}_1(\lambda)\\ \vdots \\ \mathcal{F}_r(\lambda)\end{pmatrix}^T \theta^{m-1}u.$$

Thus, $\limsup_{m\to\infty}(\mathbb{E}\mathcal{H}_m^{\mathrm{fin}})^{1/m}\leqslant\theta$. Similarly, one can show that $\lim_{m\to\infty}\hat{S}(m)^{1/m}=\varrho$ by taking eigenvectors $v_1\geqslant\mathbb{1}$ and $v_2\leqslant\mathbb{1}$. Analogously to the proofs of Lemma 4.8 and Propositions 4.9 and 4.17, we obtain the claim. □

Proof of Corollary 3.16. First, we remark that we dropped the assumption on symmetry of the laws μ_i in the case of free products of finite groups. This assumption is needed in the general case to ensure $F(e,x|\lambda)<1$. This inequality holds also in the present setting: by Woess [**25**, Equation (9.20)],

$$\alpha_i z G(e,e|z) = G_i(e_i,e_i|\xi_i(z))\xi_i(z).$$

Since $G(e,e|R)<\infty$ and $G_i(e_i,e_i|1)=\infty$, we must have $\xi_i(R)<1$, and consequently,

$$F(e,x_1\dots x_k|\lambda) = \prod_{j=1}^k F_{\tau(x_j)}(e_{\tau(x_j)},x_j|\xi_{\tau(x_j)}(\lambda)) < \prod_{j=1}^k F_{\tau(x_j)}(e_{\tau(x_j)},x_j|1) = 1.$$

In order to show that $-\log\theta/\log\alpha$ is a lower bound for $\mathrm{HD}^{\mathrm{fin}}(\Lambda)$, we can follow the reasoning in [**13**, Section 6] or also as in Section 4.2.2. Analogously to the proof of Theorem 3.8, we obtain that $\mathrm{HD}^{\mathrm{fin}}(\Omega)=\mathrm{BD}^{\mathrm{fin}}(\Omega)$. □

4.6. *Proof of Corollary* 3.18

First, we prove the following lemma:

LEMMA 4.23.
$$\overline{\mathrm{BD}^{(H)}}(\Lambda) \leqslant -\frac{\log \theta_H}{\log \alpha} \quad \textit{and} \quad \mathrm{BD}^{(H)}(\Omega) = -\frac{\log \varrho_H}{\log \alpha}.$$

Proof. First, we define the matrices $N_0 = (n_0(i,j))_{i,j\in\mathcal{I}}$ and $D_{0,H} = (d_{0,H}(i,j))_{i,j\in\mathcal{I}}$ by
$$n_0(i,j) := \begin{cases} \mathcal{F}_i^{(H)}(\lambda) & \text{if } i=j, \\ 0, & \text{otherwise,} \end{cases} \quad \text{and} \quad d_{0,H}(i,j) := \begin{cases} [\Gamma_i : H_i] - 1 & \text{if } i=j, \\ 0 & \text{otherwise.} \end{cases}$$
For $m \in \mathbb{N}$, we denote by $\mathcal{H}_m^{(H)}$ the set of words of the form $g_1 \dots g_m h \in \Gamma$ in the sense of (3.3). Since every path from e to $g_1 \dots g_m h \in \Gamma$ has to pass through points $g_1 \dots g_j h_j \in \Gamma$, where $h_j \in H$ with $h_m = h$, we have
$$\sum_{g_1\dots g_m h \in \Gamma} F_H(g_1 \dots g_m h|z) = \sum_{g_1\dots g_m h\in\Gamma} \sum_{h_1,\dots,h_{m-1}\in H} \prod_{i=1}^m F_H(g_i h_i|z) = \mathbb{1}^T N_0 N^{m-1}\mathbb{1}.$$
Choose now an eigenvector $v = (v_1, \dots, v_r)^T \geqslant \mathbb{1}$ with respect to the eigenvalue θ_H of N. Then
$$\mathbb{E}|\mathcal{H}_m^{(H)}| \leqslant \mathbb{1}^T N_0 N^{m-1}\mathbb{1} \leqslant \mathbb{1}^T N_0 N^{m-1} v = \theta_H^{m-1} \cdot \left(\sum_{i\in\mathcal{I}} v_i \mathcal{F}_i^{(H)}(\lambda)\right),$$
and therefore, $\limsup_{m\to\infty} \mathbb{E}|\mathcal{H}_m^{(H)}|^{1/m} \leqslant \theta_H$. Furthermore, we remark that $\hat{S}_H(m) = |\{x_1 \dots x_m \mid x_i \in \bigcup_{j\in\mathcal{I}} \mathcal{R}_j \setminus \{e_j\}, x_i \in \mathcal{R}_j \Rightarrow x_{i+1} \notin \mathcal{R}_j\}|$ can be written as
$$\hat{S}_H(m) = \mathbb{1}^T D_{0,H} D_H^{m-1}\mathbb{1}.$$
Taking eigenvectors $v_1 \geqslant \mathbb{1}$ and $v_2 \leqslant \mathbb{1}$ with respect to ϱ_H leads to $\lim_{m\to\infty} |\hat{S}_H(m)|^{1/m} = \varrho_H$. The same reasoning as used in the proofs of Lemma 4.8 and Propositions 4.9 and 4.17 yields the proposed claim. □

Proof of Corollary 3.18. It is sufficient to show that $-\log\theta_H/\log\alpha$ is also a lower bound for $\mathrm{HD}^{(H)}(\Lambda)$. First, we remark that, for $m \in \mathbb{N}$,
$$\begin{aligned} \sum_{g_1\dots g_m h\in\Gamma : g_1\notin\mathcal{R}_1} \mathbb{E}Z_\infty(g_1\dots g_m h) &= \sum_{g_1\dots g_m h\in\Gamma : g_1\notin\mathcal{R}_1} F(e, g_1\dots g_m h|\lambda) \\ &= \sum_{g_1\dots g_m h\in\Gamma : g_1\notin\mathcal{R}_1} \sum_{h_0\in H} F_H(g_1\dots g_m h_0|\lambda)\, F(e, h_0^{-1}h|\lambda). \end{aligned}$$
Since $|H| < \infty$, there are constants $d, D > 0$ such that $d \leqslant F(e,h|\lambda) \leqslant D$, for all $h \in H$. We write $\mathbb{1}_0 := (0,1,\dots,1)^T \in \mathbb{R}^r$ and obtain:
$$\left(\sum_{g_1\dots g_m h\in\Gamma : g_1\notin\mathcal{R}_1} \mathbb{E}Z_\infty(g_1\dots g_m h)\right)^{1/m} \leqslant (D \cdot \mathbb{1}_0^T N_0 N^{m-1}\mathbb{1})^{1/m} \xrightarrow{m\to\infty} \theta_H$$
and
$$\left(\sum_{g_1\dots g_m h\in\Gamma : g_1\notin\mathcal{R}_1} \mathbb{E}Z_\infty(g_1\dots g_m h)\right)^{1/m} \geqslant (d \cdot \mathbb{1}_0^T N_0 N^{m-1}\mathbb{1})^{1/m} \xrightarrow{m\to\infty} \theta_H.$$

This can easily be verified by substituting $\mathbb{1}$ by an eigenvector $v_1 \geqslant \mathbb{1}$ of θ_H, by an eigenvector $v_2 \leqslant \mathbb{1}$ of θ_H, respectively. With the help of this convergence behaviour and the last lemma, we can prove once again analogously to the reasoning in [**13**, Section 6] or Section 4.2.2 that the upper bounds in Lemma 4.23 equal the Hausdorff and the Box-Counting dimensions. Analogously to the proof of Theorem 3.8, we obtain that $\mathrm{HD}^{(H)}(\Omega) = \mathrm{BD}^{(H)}(\Omega)$. □

Acknowledgement. The authors are grateful to the referee whose comments led to an essential improvement of the paper. The authors are also grateful to Steven Lalley for discussions during his visit at Graz University of Technology.

References

1. I. Benjamini and S. Müller, 'On the trace of branching random walks', *Groups, Geometry, and Dynamics*, Preprint, 2010, arxiv.org/abs/1002.2781v2.
2. I. Benjamini and Y. Peres, 'Markov chains indexed by trees', *Ann. Probab.* 22 (1994) 219–243.
3. V. Cammarota and E. Orsingher, 'Cascades of particles moving at finite velocity in hyperbolic spaces', *J. Stat. Phys.* 133 (2008) 1137–1159.
4. E. Candellero and L. Gilch, 'Phase transitions for random walk asymptotics on free products of groups', *Random Structures Algorithms*, Preprint, 2009, arxiv.org/abs/0909.1893.
5. D. Cartwright and P. Soardi, 'Random walks on free products, quotients and amalgams', *Nagoya Math. J.* 102 (1986) 163–180.
6. N. Gantert and S. Müller, The critical branching Markov chain is transient', *Markov Process. Related Fields* 12 (2007) 805–814.
7. P. Gerl and W. Woess, 'Local limits and harmonic functions for nonisotropic random walks on free groups', *Probab. Theory Related Fields* 71 (1986) 341–355.
8. L. A. Gilch, 'Rate of escape of random walks', PhD Thesis, University of Technology Graz, Austria, 2007.
9. L. A. Gilch, 'Rate of escape of random walks on free products', *J. Aust. Math. Soc.* 83 (2007) 31–54.
10. L. A. Gilch, 'Asymptotic entropy of random walks on free products', *Electron. J. Probab.* 16 (2011) 76–105.
11. A. Grigor'yan and M. Kelbert, 'Recurrence and transience of branching diffusion processes on Riemannian manifolds', *Ann. Probab.* 31 (2003) 244–284.
12. T. E. Harris, *The theory of branching processes*, Die Grundlehren der mathematischen Wissenschaften, Bd. 119 (Springer, Berlin, 1963).
13. I. Hueter and S. Lalley, 'Anisotropic branching random walks on homogeneous trees', *Probab. Theory Related Fields* 116 (2000) 57–88.
14. F. I. Karpelevich, E. A. Pechersky and Y. M. Suhov, 'A phase transition for hyperbolic branching processes', *Comm. Math. Phys.* 195 (1998) 627–642.
15. S. Lalley, 'Finite range random walk on free groups and homogeneous trees', *Ann. Probab.* 21 (1993) 2087–2130.
16. S. Lalley, 'Random walks on infinite free products and infinite algebraic systems of generating functions', Preprint, http://www.stat.uchicago.edu/~lalley/Papers/RWFP.pdf.
17. S. P. Lalley and T. Sellke, 'Limit set of a weakly supercritical contact process on a homogeneous tree', *Ann. Probab.* 26 (1996) 644–657.
18. S. P. Lalley and T. Sellke, 'Hyperbolic branching brownian motion', *Probab. Theory Related Fields* 108 (1997) 171–192.
19. J. Mairesse and F. Mathéus, 'Random walks on free products of cyclic groups', *J. London Math. Soc.* 75 (2007) 47–66.
20. J. McLaughlin, 'Random walks and convolution operators on free products', PhD Thesis, New York University, 1986.
21. M. V. Menshikov and S. E. Volkov, 'Branching Markov chains: Qualitative characteristics', *Markov Process. Related Fields.* 3 (1997) 225–241.
22. S. Sawyer, 'Isotropic random walks in a tree', *Z. Wahr. Verw. Geb.* 42 (1978) 279–292.
23. D. Voiculescu, 'Addition of certain non-commuting random variables', *J. Funct. Anal.* 66 (1986) 323–346.
24. W. Woess, 'Nearest neighbour random walks on free products of discrete groups', *Boll. Un. Mat. Ital.* 5-B (1986) 961–982.
25. W. Woess, *Random walks on infinite graphs and groups* (Cambridge University Press, Cambridge, 2000).

Elisabetta Candellero
Institut für mathematische Strukturtheorie (Math. C)
Graz University of Technology
Steyrergasse 30
A-8010 Graz
Austria

candellero@TUGraz.at
http://www.math.tugraz.at/∼candellero/

Lorenz Gilch
Section de Mathématiques
Université de Genève
2-4, rue du Lièvre
1211 Genève 4
Switzerland

gilch@TUGraz.at
http://www.math.tugraz.at/∼gilch/

Sebastian Müller
LATP
Université de Provence
39, rue F. Joliot Curie
13453 Marseille
France

mueller@cmi.univ-mrs.fr
http://www.latp.univ-mrs.fr/∼mueller

Publication D

Asymptotic Entropy of Random Walks on Regular Languages over a Finite Alphabet

L. Gilch

Accepted for publication (modulo minor revision) in *Electronic Journal of Probability*, 2015.

ASYMPTOTIC ENTROPY OF RANDOM WALKS ON REGULAR LANGUAGES OVER A FINITE ALPHABET

LORENZ A. GILCH

Abstract. We prove existence of asymptotic entropy of random walks on regular languages over a finite alphabet and we give formulas for it. Furthermore, we show that the entropy varies real-analytically in terms of probability measures of constant support, which describe the random walk. This setting applies, in particular, to random walks on virtually free groups.

1. Introduction

Let $\mathcal{A}$ be a finite alphabet and let $\mathcal{A}^*$ be the set of all finite words over the alphabet $\mathcal{A}$, where o denotes the empty word. Consider a transient Markov chain $(X_n)_{n\in\mathbb{N}_0}$ on $\mathcal{A}^*$ with $X_0 = o$ such that at each instant of time the last $K \in \mathbb{N}$ letters of the current word may be replaced by a word of length of at most $2K$ and the transition probabilities depend only on the last K letters of the current word and on the replacing word. For better visualization and ease of presentation, we also consider the random walk on $\mathcal{A}^*$ as a random walk on an undirected graph $\mathcal{G}$. Denote by π_n the distribution of X_n. We are interested whether the sequence $\frac{1}{n}\mathbb{E}[-\log \pi_n(X_n)]$ converges, and if so to describe the limit. If it exists, it is called the *asymptotic entropy*, which was introduced by Avez [1]. The aim of this paper is to prove existence of the asymptotic entropy, to describe it as the rate of escape w.r.t. the Greenian distance and to prove its real-analytic behaviour when varying the transition probabilities of constant support.

We outline some background on this topic. Random Walks on regular languages have been studied by e.g. Lalley [16] and Malyshev [19] amongst others. Concerning asymptotic entropy it is well-known by Kingman's subadditive ergodic theorem (see Kingman [15]) that the entropy exists for random walks on groups if $\mathbb{E}[-\log \pi_1(X_1)] < \infty$. In contrast to this fact existence of the entropy on more general structures is not known a priori. In our setting we are not able to apply the subadditive ergodic theorem since we neither have subadditivity nor a global composition law of words if the random walk is performed on a proper subset of $\mathcal{A}^*$ (that is, not every word $w \in \mathcal{A}^*$ can be reached from o with positive probability). This forces us to use other techniques like generating functions techniques. These generating functions are power series with probabilities as coefficients, which describe the characteristic behaviour of the underlying random walks. The technique of our proof

Date: August 19, 2015.
2000 *Mathematics Subject Classification.* Primary: 60J10; Secondary: 28D20.
Key words and phrases. random walks, regular languages, entropy, analytic.

of existence of the entropy was motivated by Benjamini and Peres [2], where it is shown that for random walks on groups the entropy equals the rate of escape w.r.t. the Greenian distance; compare also with Blachère, Haïssinsky and Mathieu [3]. In particular, we will also show that the asymptotic entropy h is the rate of escape w.r.t. a distance function in terms of Green functions, which in turn yields that h is also the rate of escape w.r.t. the Greenian distance. Moreover, we prove convergence in probability and convergence in L_1 of the sequence $-\frac{1}{n}\log \pi_n(X_n)$ to h, and we show also that h can be computed along almost every sample path as the limes inferior of the aforementioned sequence. The question of almost sure convergence of $-\frac{1}{n}\log \pi_n(X_n)$ to some constant h, however, remains open. Similar results concerning existence and formulas for the entropy are proved in Gilch and Müller [9] for random walks on directed covers of graphs and in Gilch [8] for random walks on free products of graphs. Furthermore, we give formulas for the entropy which allow numerical computations and also exact calculations in some special cases. The main idea in our proofs is to fix a priori a sequence of nested cones in the associated graph $\mathcal{G}$ and to track the random walk's way to infinity through these cones. Similar ideas have been used independently by Woess [23] for context-free pairs of groups. The techniques in our proofs are restricted to the case of bounded range random walks: in the case of unbounded range the situation gets much more complicated since Martin and Gromov boundaries may differ even under assumption of some exponential moments to be finite; compare with Gouëzel [10].

Kaimanovich and Erschler asked whether drift and entropy of random walks vary continuously (or even analytically) when varying the probabilities of the random walk with keeping the support of single step transitions constantly. In view of this question we also show in this article that h is real-analytic in terms of the parameters describing the random walk on $\mathcal{A}^*$. This fact applies, in particular, to the case of bounded range random walks on virtually free groups, which goes beyond the scope of previous results related to the question of analyticity. Ledrappier [17] showed that the entropy varies real-analytically for finitely supported random walks on free groups; with the help of "barriers" (that is, nested sequences of subsets which have to be passed successively) and the study of Martin kernels he identifies the entropy as the boundary entropy. The present article uses also some kind of barriers (called "cones") to track the random walk's path to infinity, but the approach is different: here, we identify the entropy as the Shannon entropy of a hidden Markov chain (see Theorem 2.5), which arises from splitting up the random walk into pieces between the entries of these nested cones. For some special cases (e.g., free groups) we even give a formula (see Theorem 7.4) for the entropy of the hidden Markov chain, which allows numerical calculations. A similar idea for proving existence of the entropy has also been used in Gilch [8] for random walks on free products of graphs by cutting the random walk into pieces; Theorem 7.4 applies also to free products of *finite* graphs, but not necessarily for free products of infinite graphs. The important difference between [8] and the present article is that analyticity of the entropy in [8] follows directly from the formulas for the entropy, while we have to make much more effort to show this property in the present context of regular languages. Finally, let us remark that random walks on regular languages do not only extend results from free groups or free products to the next general case like virtually free groups but also to a wider class like context-free graphs (see Subsection 2.2).

At this point let us summarize further papers concerning continuity and analyticity of the drift and entropy that have been published recently: Ledrappier [18] showed that the drift and entropy of finitely supported random walks on hyperbolic groups are Lipschitz, while Mathieu [20] showed that the entropy of symmetric, finitely supported random walks in hyperbolic groups are differentiable; Haïssinsky, Mathieu and Müller [11] proved analyticity of the drift for random walks on surface groups. The recent survey article of Gilch and Ledrappier [6] collects several results about analyticity of drift and entropy of random walks on groups.

The basic reasoning of our proofs follows a similar argumentation as in [9] and [8], but since a straight-forward adaption is not possible we have to do more effort in the present setting: we will show that the entropy equals the rate of escape w.r.t. some special length function, and we deduce the proposed properties analogously. For the proof of analyticity of the entropy we will extract a hidden Markov chain from our random walk and we will apply a result of Han and Marcus [12]. The plan of the paper is as follows: in Sections 2 and 3 we define the random walk on $\mathcal{A}^*$ and the associated generating functions. Section 4 explains the construction of cones in the present context. In Sections 5 and 6 we prove existence of the asymptotic entropy and give a formula for it, while in Section 7 we give estimates and a more explicit formula in some special case. Section 8 shows real-analyticity of the entropy.

2. Random Walks on Regular Languages

2.1. **Definitions and Main Results.** Let $\mathcal{A}$ be a finite alphabet and denote by $\mathcal{A}^*$ the set of all finite words over $\mathcal{A}$. We write o for the empty word and $\mathcal{A}^n$, $n \in \mathbb{N}$, for the set of all words over $\mathcal{A}$ consisting of exactly n letters. For two words $w_1, w_2 \in \mathcal{A}^*$, $w_1 w_2$ denotes the concatenated word. A *random walk on a regular language* is a Markov chain $(X_n)_{n\in\mathbb{N}_0}$ on the set $\mathcal{A}^* = \bigcup_{n\geq 1} \mathcal{A}^n \cup \{o\}$, whose transition probabilities obey the following rules:

(i) Only the last two letters of the current word may be modified.
(ii) Only one letter may be adjoined or deleted at one instant of time.
(iii) Adjunction and deletion may only be done at the end of the current word.
(iv) Probabilities of modification, adjunction or deletion depend only on the last two letters of the current word and on the substitute letters.

Compare with Lalley [16] and Gilch [7]. In other words, at each step the last two letters of the current word may be replaced by a non-empty word of length of at most 3 and the transition probabilities depend only on the last two letters of the current word and the replacing word of length of at most 3. More formally, the transition probabilities of the

Markov chain $(X_n)_{n\in\mathbb{N}_0}$ can be written as follows, where $w \in \mathcal{A}^*$, $a_1, a_2, b_1, b_2, b_3 \in \mathcal{A}$:

$$\begin{aligned}
\mathbb{P}[X_{n+1} = wb_1b_2 \mid X_n = wa_1a_2] &= p(a_1a_2, b_1b_2),\\
\mathbb{P}[X_{n+1} = wb_1b_2b_3 \mid X_n = wa_1a_2] &= p(a_1a_2, b_1b_2b_3),\\
\mathbb{P}[X_{n+1} = wb_1 \mid X_n = wa_1a_2] &= p(a_1a_2, b_1),\\
\mathbb{P}[X_{n+1} = b_1 \mid X_n = a_1] &= p(a_1, b_1), \qquad (2.1)\\
\mathbb{P}[X_{n+1} = b_1b_2 \mid X_n = a_1] &= p(a_1, b_1b_2),\\
\mathbb{P}[X_{n+1} = o \mid X_n = a_1] &= p(a_1, o),\\
\mathbb{P}[X_{n+1} = b_1 \mid X_n = o] &= p(o, b_1),\\
\mathbb{P}[X_{n+1} = o \mid X_n = o] &= p(o, o).
\end{aligned}$$

Not all of these probabilities need to be strictly positive. Initially, we set $X_0 := o$. If we start the random walk at $w \in \mathcal{A}^*$ instead of o, we write $\mathbb{P}_w[\,\cdot\,] := \mathbb{P}[\,\cdot \mid X_0 = w]$. For $w_1, w_2 \in \mathcal{A}^*$, the n-step transition probabilities are denoted by $p^{(n)}(w_1, w_2) := \mathbb{P}_{w_1}[X_n = w_2]$. The set of *accessible words* from o is given by

$$\mathcal{L} = \bigl\{w \in \mathcal{A}^* \mid \exists n \in \mathbb{N} : \mathbb{P}[X_n = w \mid X_0 = o] > 0\bigr\}.$$

We will also think of the random walk $(X_n)_{n\in\mathbb{N}_0}$ as a nearest neighbour random walk on an *undirected graph* $\mathcal{G}$, where the vertices are the elements of $\mathcal{L}$ and undirected edges are between two vertices if and only if one can walk from one word to the other one in a single step. For this purpose, we need the following assumption:

Assumption 2.1 (Weak symmetry)**.** *For all $u, v \in \mathcal{A}^*$ we assume that $\mathbb{P}_u[X_1 = v] > 0$ implies $\mathbb{P}_v[X_1 = u] > 0$. We call this property weak symmetry.*

In particular, Assumption 2.1 yields irreducibility of the random walk on $\mathcal{L}$. Moreover, this assumption will be necessary for the construction of a sequence of cones in the graph $\mathcal{G}$ which track the random walk's way to infinity. As the interested reader will see, weak symmetry can obviously be weakened in some way but for reason of better readability we keep this natural assumption; for a discussion on this assumption, we refer to Appendix A.2.

Since the purpose of the paper is the investigation of the asymptotic behaviour of transient random walks, we obviously need that $\mathcal{L}$ is infinite in our setting. It is an easy exercise to check that the set $\mathcal{L}$ is a *regular language* over the alphabet $\mathcal{A}$, that is, the words are accepted by a finite-state automaton. For more details on regular languages, we refer e.g. to Hopcraft and Ullman [13]. Since we make no further use of the theory of languages, we will not discuss this in more detail but we remark the recursive structure of regular languages. Let us note that bounded range random walks on *virtually free groups* constitute a special case of our setting, and our results directly apply; see e.g. Lalley [16]. Thus, our results apply directly to a large class of random walks on groups and go beyond recent results for random walks on groups.

Remark 2.2. *Observe that the assumption that transition probabilities depend only on the* last two *letters of the current word and that changes of the current word involve only the* last two *letters may be weakened to dependence and changes of the last $K \in \mathbb{N}$ letters* *of the current word and replacements of the last K letters by words of length of at most $2K$. This is done by blocking words of length of at most K to new single letters; see* [16, Section 3.3] *for further details and comments. If we make further assumptions on our random walk in the following, we will show that it does not depend on the fact if we use the "blocked letter language" (that is, dependence on the last two letters as given by (2.1) after an application of the "recoding trick") or the general case (dependence on the last K letters as given by (B.1)), that is, no required properties are lost when switching from the K-dependent case to the "blocked letter language"; for further comments, see Appendix B. It will turn out that the K-dependent case works completely analogously as the "blocked letter language" case; however, the derived equations and formulas are much more complex, so we restrict ourselves onto the case where the random walk is defined as at the beginning of this section via (2.1). In particular, there is no additional gain in the techniques and proofs when investigating the K-dependent case. Finally, let us note that it is not sufficient to consider the case where the transition probabilities/changes of words involve only the last letter in order to be able to apply this recoding trick!*

We introduce some notation. The *natural word length* of any $w \in \mathcal{A}^*$ is denoted by $|w|$. If $w \in \mathcal{A}^*$ and $k \in \mathbb{N}$ with $|w| \geq k$ then $w[k]$ denotes the k-th letter of w, and $[w]$ denotes the last two letters of w when $w \neq o$ is not a single letter.

Malyshev [19] proved that the rate of escape w.r.t. the natural word length exists for irreducible random walks on regular languages, that is, there is a non-negative constant ℓ such that

$$\lim_{n\to\infty} \frac{|X_n|}{n} = \ell \quad \text{almost surely.}$$

Here, ℓ is called the *rate of escape*. Furthermore, by [19] follows that ℓ is strictly positive if and only if $(X_n)_{n\in\mathbb{N}_0}$ is transient. In [7] there are explicit formulas for the rate of escape w.r.t. more general length functions.

Another characteristic number of random walks is the asymptotic entropy. Denote by π_n the distribution of X_n. If there is a non-negative constant h such that the limit

$$h = \lim_{n\to\infty} -\frac{1}{n}\mathbb{E}\bigl[\log \pi_n(X_n)\bigr]$$

exists, then h is called the *asymptotic entropy*. Since we only have a partial composition law for concatenation of two words (if $\mathcal{L} \subset \mathcal{A}^*$) and since we have no subadditivity and transitivity of the random walk, we are not able to apply – as in the case of random walks on groups – Kingman's subadditive ergodic theorem in order to show existence of h. It is, however, easy to see that the entropy equals zero if the random walk is recurrent (see Corollary 7.2). Therefore, from now on we will only consider *transient* random walks $(X_n)_{n\in\mathbb{N}_0}$.

Remark 2.3. *Observe that $\lim_{n\to\infty} -\frac{1}{n}\log \pi_n(X_n)$ is not necessarily deterministic: take two homogeneous trees of different degrees $d_1, d_2 \geq 3$; identify their root with one single root which becomes o and consider the simple random walk on this new inhomogeneous tree with starting point o. Obviously, this random walk can be modelled as a random walk on a regular language. Then the limit $\lim_{n\to\infty} -\frac{1}{n}\log \pi_n(X_n)$ depends on the fact in which of the two subtrees the random walks goes to infinity. Hence, the sequence $-\frac{1}{n}\log \pi_n(X_n)$*

converges with probability $d_1/(d_1+d_2)$ *to* $\log(d_1-1)$ *and with probability* $d_2/(d_1+d_2)$ *to* $\log(d_2-1)$*; this can, e.g., be calculated by the formulas given in* [8].

We have to make another assumption on the transition probabilities:

Assumption 2.4 (Suffix-irreducibility)**.** *We assume that the random walk on* $\mathcal{L}$ *is suffix-irreducible, that is, for all* $w = w_0a_0b_0 \in \mathcal{L}$ *with* $w_0 \in \mathcal{A}^*, a_0b_0 \in \mathcal{A}^2$ *and for all* $ab \in \mathcal{A}^2$ *there is* $n \in \mathbb{N}$ *and* $w_1 \in \mathcal{A}^*$ *such that*

$$\mathbb{P}\Big[X_n = w_0w_1ab, \forall k \leq n : |X_k| \geq |w| \,\Big|\, X_0 = w\Big] > 0.$$

This assumption excludes degenerate cases and will guarantee existence of ℓ; compare with [7, End of Section 2.1]. We remark that famous previous papers about random walks on regular languages (in particular, the basic ones of [19] and [16]) require stronger assumptions than this non-degeneracy assumption. Later on it will be clear that one can relax this condition in some way without needing additional techniques or ideas for the proofs. Hence, for purpose of ease and better readability, we keep this assumption until further notice. We will give further comments on this assumption in Appendix A.1.

The main idea behind our proofs will be the construction of an a priori fixed sequence of cones (that is, special subsets of $\mathcal{L}$), from which we extract a subsequence of nested cones which gives the information how the random walk tends to infinity. This extraction will be done via a hidden Markov chain $(\mathbf{Y}_k)_{k\in\mathbb{N}}$ with an underlying positive recurrent Markov chain: the asymptotic entropy $H(\mathbf{Y})$ of the process $(\mathbf{Y}_k)_{k\in\mathbb{N}}$ is given by (5.6). The average distance between two nested cones will be denoted by λ which is given by (5.8): if $X_{\mathbf{e}_k}$ denotes the word (i.e., the vertex in $\mathcal{G}$) where the k-th nested subcone is finally entered with no further exits of this cone, then $\lambda = \mathbb{E}[|X_{\mathbf{e}_2}| - |X_{\mathbf{e}_1}|]$. Our first main result concerns existence of the asymptotic entropy, which is finally proven in Section 6:

Theorem 2.5. *Consider a transient random walk* $(X_n)_{n\in\mathbb{N}_0}$ *on a regular language, which satisfies Assumptions 2.1 and 2.4. Then the asymptotic entropy* h *of* $(X_n)_{n\in\mathbb{N}_0}$ *exists and equals*

$$h = \frac{\ell \cdot H(\mathbf{Y})}{\lambda},$$

where $H(\mathbf{Y})$ *is given by (5.6) and* λ *by (5.8).*

Recall that the random walk is described by the values in (2.1). A natural question is whether the entropy varies regularly if the parameters in (2.1) are varied slightly and if positive transition probabilities remain positive by this variation. The following result gives an answer to this question, where the proof is given in Section 8:

Theorem 2.6. *For transient random walks on regular languages satisfying Assumptions 2.1 and 2.4, the entropy* h *varies real-analytically under all probability measures of constant support.*

Moreover, we can also describe the asymptotic entropy in the following way:

Corollary 2.7. *We have the following types of convergence:*

(1) *For almost every trajectory of the random walk* $(X_n)_{n\in\mathbb{N}_0}$*,*

$$h = \liminf_{n\to\infty} -\frac{1}{n}\log \pi_n(X_n).$$

(2) *Convergence in probability:*

$$-\frac{1}{n}\log \pi_n(X_n) \xrightarrow{\mathbb{P}} h.$$

(3) *Convergence in* L_1*:*

$$-\frac{1}{n}\log \pi_n(X_n) \xrightarrow{L_1} h.$$

The *Greenian distance* between two words $w_1, w_2 \in \mathcal{L}$ is defined as

$$d_{\mathrm{Green}}(w_1,w_2) := -\log \mathbb{P}[\exists n \in \mathbb{N}_0 : X_n = w_2 \mid X_0 = w_1].$$

Analogously to the situation for random walks on groups, we get the following result, which is finally proven at the end of Section 6:

Corollary 2.8. *The entropy is the rate of escape with respect to the Greenian distance, that is,*

$$h = \lim_{n\to\infty} -\frac{1}{n} d_{\mathrm{Green}}(o, X_n) \quad \textit{almost surely.}$$

Further results are given in Section 7, where we show that $h > 0$ (Corollary 7.1) for non-degenerate transient random walks, give an inequality between entropy, drift and growth (Theorem 7.3) and give an exact formula in some special case (Theorem 7.4).

2.2. **Examples.** We give three classical examples for regular languages.

2.2.1. *Stacks.* In computer science theory *stacks* play an important role for modelling algorithms. In this setting letters represent different procedures and words are lists of procedures, which are called randomly. The last letter of the current word is the actual running procedure which may produce more subprocedures or will finish some open procedures, which in turn yields that the stack is getting larger or smaller randomly. Thus, this setting can be encoded by regular languages. Compare also with Lalley [16].

2.2.2. *Virtually Free Groups.* An important class of examples is given by *virtually free groups*, that is, groups which contain a free group as a subgroup of finite index. Let Γ be a virtually free group which contains the free group $\mathbb{F}_d$ with d generators as a subgroup of index $[\Gamma : \mathbb{F}_d] = k$. Let $\mathbb{F}_d$ be generated by the elements $a_1, a_1^{-1}, \dots, a_d, a_d^{-1}$, and let $h_1, \dots, h_k$ be representants of the k different left co-sets of Γ. That is, each element $x \in \Gamma$ can be written as

$$x_1 x_2 \dots x_m h_j,$$

where $m \in \mathbb{N}_0$, $j \in \{1, \dots, k\}$ and $x_1, \dots, x_m \in \{a_1, a_1^{-1}, \dots, a_d, a_d^{-1}\}$ such that $x_i^{-1} \neq x_{i+1}$ for all $i \in \{1, \dots, m-1\}$. Now it is clear that each group invariant, finitely supported random walk on Γ can be considered as a random walk on a regular language with alphabet

$\mathcal{A} = \{a_1, a_1^{-1}, \dots, a_d, a_d^{-1}, h_1, \dots, h_k\}$ since multiplication from the right changes only a bounded number of letters at the end of the current word. Compare also with the detailed example of free products with amalgamation in [7, Section 3.1]

2.2.3. *Context-Free Graphs.* Another important class is given by *context-free graphs*, and in particular by certain *Schreier graphs*, which can also be considered as random walks on regular languages. This class justifies the study of random walks on regular languages in its own right and not only as an extension of free groups or free products. We sketch the concept of context-free graphs: consider a labelled, symmetric graph $\mathcal{G}$ with root $\mathbf{r}$. Consider the connected components of $\mathcal{G}$ after removing all vertices (and adjoint edges) which are at distance less or equal than some $n \in \mathbb{N}$ to $\mathbf{r}$. If there are only finitely many different isomorphism types as labelled graphs of these connected components then the graph is called *context-free*; see Muller and Schupp [22]. We give a short explanation why these graphs fit into the setting of regular languages: later the mindful reader will notice that our random walks are performed on some graph with finitely many different cone types (that is, finitely many different isomorphism classes of connected components after removal of all vertices at distance less or equal than n to $\mathbf{r}$). Since there are only finitely many different cone types one can deduce a finite-state automaton from the context-free graph, which accepts just the words which describe the different vertices of $\mathcal{G}$. As a specific example, consider a virtually free group, a finitely generated free subgroup and an associated Schreier graph: by Woess [23, Theorem 2.10], the Schreier graph satisfies all needed irreducibility requirements. For further details, we refer to Muller and Schupp [21], [22] and Ceccherini-Silberstein and Woess [4] and [23].

3. Generating Functions

For $w_1, w_2 \in \mathcal{A}^*$, $z \in \mathbb{C}$, the *Green function* is defined as

$$G(w_1, w_2|z) := \sum_{n\geq 0} p^{(n)}(w_1, w_2) \cdot z^n$$

and the *last visit generating function* as

$$L(w_1, w_2|z) := \sum_{n\geq 0} \mathbb{P}\bigl[X_n = w_2, \forall m \in \{1, \dots, n\} : X_m \neq w_1 \bigm| X_0 = w_1\bigr] \cdot z^n.$$

By conditioning on the last visit to w_1, an important relation between these functions is given by

$$G(w_1, w_2|z) = G(w_1, w_1|z) \cdot L(w_1, w_2|z). \tag{3.1}$$

In the following we introduce further generating functions, which also have been used analogously in [7]. Define for $a, b, c, d, e \in \mathcal{A}$ and real $z > 0$

$$H(ab, c|z) := \sum_{n\geq 1} \mathbb{P}\bigl[X_n = c, \forall m < n : |X_m| > 1 \bigm| X_0 = ab\bigr] \cdot z^n$$

and

$$\bar{L}(ab,cde|z) := \sum_{n\geq 1} \mathbb{P}\bigl[X_n = cde, |X_{n-1}| = 2, \forall m \in \{1,\dots,n\} : |X_m| \geq 2, \bigl| X_0 = ab\bigr] \cdot z^n,$$

$$\overline{G}(ab,cd|z) := \sum_{n\geq 0} \mathbb{P}\bigl[X_n = cd, \forall m \in \{1,\dots,n\} : |X_m| \geq 2 \bigl| X_0 = ab\bigr] \cdot z^n.$$

We write $\bar{L}(ab,cde) := \bar{L}(ab,cde|1)$. These generating functions can be computed in two steps: first, one solves the following system of equations which arises by case distinction on the first step:

$$\begin{aligned} H(ab,c|z) &= p(ab,c)\cdot z + \sum_{de\in\mathcal{A}^2} p(ab,de)\cdot z\cdot H(de,c|z) \\ &\quad + \sum_{def\in\mathcal{A}^3} p(ab,def)\cdot z\cdot \sum_{g\in\mathcal{A}} H(ef,g|z)\cdot H(dg,c|z); \end{aligned} \tag{3.2}$$

compare with [16] and [7]. The system (3.2) consists of equations of quadratic order, and therefore the functions $H(\cdot,\cdot|z)$ are algebraic, if the transition probabilities are algebraic. We now get the functions $\overline{G}(ab,cd|z)$ by solving the following linear system of equations which also arises by case distinction on the first step:

$$\begin{aligned} \overline{G}(ab,cd|z) &= \delta_{ab}(cd) + \sum_{c_1d_1\in\mathcal{A}^2} p(ab,c_1d_1)\cdot z\cdot \overline{G}(c_1d_1,cd|z) + \\ &\quad + \sum_{c_1d_1e_1\in\mathcal{A}^3} p(ab,c_1d_1e_1)\cdot z\cdot \sum_{f\in\mathcal{A}} H(d_1e_1,f|z)\cdot \overline{G}(c_1f,cd|z). \end{aligned}$$

Finally, we get

$$\bar{L}(ab,cde|z) = \sum_{a_1b_1\in\mathcal{A}^2} \overline{G}(ab,a_1b_1|z)\cdot z\cdot p(a_1b_1,cde). \tag{3.3}$$

Obviously, it is sufficient to consider only those functions $H(ab,\cdot|z)$, $\overline{G}(ab,\cdot|z)$ and $L(ab,\cdot|z)$ such that there exists some $w_0 \in \mathcal{A}^*$ with $w_0ab \in \mathcal{L}$; the remaining functions do not play a role for our random walk. Moreover, one can compute the Green functions of the form $G(o,w|z)$, $w \in \mathcal{L}$ with $|w| \leq 3$, by solving

$$\begin{aligned} G(w_1,w_2|z) &= \delta_{w_1}(w_2) + \sum_{w_3\in\mathcal{A}^*:|w_3|\leq 3} p(w_1,w_3)\cdot z\cdot G(w_3,w_2|z) + \\ &\quad + \mathbb{1}_3(w_1)\cdot \sum_{cde\in\mathcal{A}^3} p(w_1[2]w_1[3],cde)\cdot z\cdot \sum_{f\in\mathcal{A}} H(de,f|z)\cdot G(w_1[1]cf,w_2|z), \end{aligned}$$

where $w_1,w_2 \in \mathcal{A}^*$ with $|w_1|,|w_2| \leq 3$ and $\mathbb{1}_3(w_1) := 1$, if $|w_1| = 3$, and $\mathbb{1}_3(w_1) := 0$ otherwise.

We also define for $ab \in \mathcal{A}^2$:

$$\xi(ab) := \mathbb{P}\bigl[\forall n \geq 0 : |X_n| \geq 2 \bigm| X_0 = ab\bigr] = 1 - \sum_{f\in\mathcal{A}} H(ab,f|1).$$

When starting at a word $wab \in \mathcal{L}$, where $w \in \mathcal{A}^*$, $\xi(ab)$ is the probability that the process $(X_n)_{n\in\mathbb{N}_0}$ will not visit any words of length $|wab|-1$ or smaller. In this case the prefix w

will remain constant for the rest of the process. Observe that, for transient random walks, $\xi(ab)>0$ for all $ab\in\mathcal{A}^2$ due to Assumption 2.4. We define a "length function" on $\mathcal{L}$ by

$$l(w):=-\log L(o,w|1) \quad \text{ for } w\in\mathcal{L}. \tag{3.4}$$

For $n\geq 2$ and $a_1,\dots,a_n\in\mathcal{A}$, the functions $L(o,a_1\dots a_n|z)$ can be rewritten as

$$\sum_{b,b_0,c_0\in\mathcal{A}} L(o,b|z)\cdot z\cdot p(b,b_0c_0) \sum_{\substack{b_1,\dots,b_{n-2}\in\mathcal{A},\\ c_1,\dots,c_{n-2}\in\mathcal{A}}} \prod_{i=1}^{n-2} \bar{L}(b_{i-1}c_{i-1},a_ib_ic_i|z)\cdot \overline{G}(b_{n-2}c_{n-2},a_{n-1}a_n|z); \tag{3.5}$$

each path from o to $a_1\dots a_n$ is decomposed to the last times when the sets $\mathcal{A},\mathcal{A}^2,\dots,\mathcal{A}^{n-1}$ are visited, that is, the factor $\bar{L}(b_{i-1}c_{i-1},a_ib_ic_i|z)$ corresponds to the parts of the paths from o to $a_1\dots a_n$ between the final exits of the sets $\mathcal{A}^i$ and $\mathcal{A}^{i+1}$.

4. Cones

4.1. **Definitions of Cones and Properties.** In this section we introduce the structure of cones in our setting. A *path* in $\mathcal{A}^*$ is a sequence of words $\langle w_0,w_1,\dots,w_m\rangle$, $m\in\mathbb{N}$, in $\mathcal{A}^*$ such that $\mathbb{P}_{w_{i-1}}[X_1=w_i]>0$ for all $1\leq i\leq m$. By weak symmetry, we have that, for each such path, the reversed sequence of words $\langle w_m,w_{m-1},\dots,w_0\rangle$ is also a path. For $n\in\mathbb{N}$, define $\mathcal{A}^*_{\geq n}:=\bigl\{w\in\mathcal{A}^*\bigm| |w|\geq n\bigr\}$. For any $w_0\in\mathcal{A}^*_{\geq 2}$, we define the *cone* rooted at w_0 as

$$C(w_0):=\left\{ w\in\mathcal{A}^*_{\geq|w_0|} \;\middle|\; \begin{array}{c}\exists m\in\mathbb{N}_0\ \exists \text{ path } \langle w_0,w_1,\dots,w_{m-1},w\rangle\\ \text{with } w_1,\dots,w_{m-1}\in\mathcal{A}^*_{\geq|w_0|}\end{array}\right\}.$$

In other words, when we consider the associated graph $\mathcal{G}$ then the cone $C(w_0)$ can be viewed as the subgraph of $\mathcal{G}$ which is the connected component containing w_0 after removing all vertices $w'\in\mathcal{A}\setminus\mathcal{A}^*_{\geq|w_0|}$ and the adjacent edges to these w'. In particular, we have $w_0\in C(w_0)$. If $w_1\in C(w_0)$ then we have $C(w_1)\subseteq C(w_0)$: indeed, let be $w_2\in C(w_1)$; therefore, $|w_2|\geq|w_1|\geq|w_0|$ and there are paths $\langle w_0,w_1',\dots,w_k',w_1\rangle$ through words $w_1',\dots,w_k'\in\mathcal{A}^*_{\geq|w_0|}$ and $\langle w_1,w_1'',\dots,w_l'',w_2\rangle$ through words $w_1'',\dots,w_l''\in\mathcal{A}^*_{\geq|w_1|}\subseteq\mathcal{A}^*_{\geq|w_0|}$. Hence, there is a path $\langle w_0,w_1',\dots,w_k',w_1,w_1'',\dots,w_l'',w_2\rangle$ through words in $\mathcal{A}^*_{\geq|w_0|}$, that is, $w_2\in C(w_0)$ yielding $C(w_1)\subseteq C(w_0)$. The cone $C(w_1)$ is then called a *subcone* of $C(w_0)$.

Observe that each element $w\in C(w_0)$ has the form $w=a_1\dots a_{m-2}\bar{w}$, where $w_0=a_1\dots a_m$ with $m\geq 2$, $a_1,\dots,a_m\in\mathcal{A}$ and where $\bar{w}\in\mathcal{A}^*_{\geq 2}$: indeed, by definition each $w\in C(w_0)$ can be reached from w_0 by a path through words of length bigger or equal than $|w_0|$. Thus, the first $m-2$ letters are *not* changed along such a path.

By the suffix-irreducibility Assumption 2.4, we have the following important property for cones: let be $w\in\mathcal{A}^*$ and $ab,cd\in\mathcal{A}^2$; then the cone $C(wab)$ has a proper subcone $C(wxcd)\subset C(wab)$ with a suitable choice of $x\in\mathcal{A}^*\setminus\{o\}$.

Recall that $[w]$ denotes the last two letters of a word $w\in\mathcal{A}^*_{\geq 2}$. We say that two cones $C(w_1)$ and $C(w_2)$, $w_1,w_2\in\mathcal{A}^*$, are *isomorphic* if $C([w_1])=C([w_2])$. The following lemma explains why we call these cones "isomorphic". Since the proof of the following lemma is elementary, we omit the proof at this place and hand it in later in Appendix C.

Lemma 4.1. *Let be $w_1 = a_1 \dots a_m$, $w_2 = b_1 \dots b_n \in \mathcal{A}^*_{\geq 2}$ with $a_1, \dots, a_m, b_1, \dots, b_n \in \mathcal{A}$ such that $C(w_1)$ and $C(w_2)$ are isomorphic. Then:*

(1) *The mapping $\varphi : C(w_1) \to C(w_2)$ defined by*

$$\varphi(a_1 \dots a_{m-2}\bar{w}) = b_1 \dots b_{n-2}\bar{w} \quad \text{for } \bar{w} \in \mathcal{A}^*_{\geq 2} \text{ with } a_1 \dots a_{m-2}\bar{w} \in C(w_1)$$

is a bijection which preserves the adjacency relation, that is, $p(w', w'') > 0$ if and only if $p\big(\varphi(w'), \varphi(w'')\big) > 0$ for all $w', w'' \in C(w_1)$.

(2) *The cones are isomorphic as subgraphs of $\mathcal{G}$.*

The lemma says implicitly that the words of two isomorphic cones differ only by different prefixes. Moreover, there is a natural 1-to-1 correspondence of paths inside $C(w_1)$ and paths in an isomorphic cone $C(w_2)$ where obviously each such path in $C(w_1)$ and the corresponding path in the other isomorphic cone $C(w_2)$ have the same probability: let be $\langle w'_0, w'_1, \dots, w'_m \rangle$ a path in $C(w_1)$; then $\langle \varphi(w'_0), \varphi(w'_1), \dots, \varphi(w'_m) \rangle$ is a path in $C(w_2)$ and

$$\mathbb{P}\big[X_1 = w'_1, \dots, X_m = w'_m \big| X_0 = w'_0\big] = \mathbb{P}\big[X_1 = \varphi(w'_1), \dots, X_m = \varphi(w'_m) \big| X_0 = \varphi(w'_0)\big].$$

We remark that $C(w)$ and $C(w')$, $w, w' \in \mathcal{A}^*_{\geq 2}$, with $C([w]) \neq C([w'])$ can still be isomorphic as subgraphs of $\mathcal{G}$ but we will still distinguish them as elements of different isomorphism classes according to our definition of isomorphism of cones.

Our construction of cones ensures that different cones are either nested in each other or disjoint as the next lemma will show; the elementary proof of the next lemma is again omitted and will be handed in later in the Appendix C.

Lemma 4.2. *Let be $w_1, w_2 \in \mathcal{A}^*_{\geq 2}$. Then the cones $C(w_1)$ and $C(w_2)$ are either nested in each other, that is, $C(w_1) \subseteq C(w_2)$ or $C(w_2) \subseteq C(w_1)$, or they are disjoint, that is, $C(w_1) \cap C(w_2) = \emptyset$. If we even have $|w_1| = |w_2|$ then we have $C(w_1) = C(w_2)$ or $C(w_1) \cap C(w_2) = \emptyset$.*

At this point let us mention that the weak symmetry Assumption 2.1 is crucial here: if this assumption is dropped then two cones $C(w_1)$ and $C(w_2)$, where $w_1, w_2 \in \mathcal{A}^*_{\geq 2}$ with $|w_1| = |w_2|$ and $C(w_1) \cap C(w_2) \neq \emptyset$ may be non-isomorphic. This case makes everything much more difficult in our proofs since the property of cones from the last lemma (either nestedness or disjointness) is lost and since we want to track the random walk's way to infinity by distinguishing which of the (disjoint) cones are successively finally entered on its way to infinity. The author is however confident that one can adapt the situation if weak symmetry does not hold but this would need much more effort with loss of good readability of our proofs and no additional gain of the techniques; for further comments see Appendix A.2.

Since isomorphism of cones depends only on the last two letters of their roots, we have obviously only finitely many different isomorphism classes of cones. These isomorphism classes can be described by two-lettered words $ab \in \mathcal{A}^2$: first, for each isomorphism class of cones we fix some ab representing the class of $C(ab)$. Let $\mathcal{J} \subseteq \mathcal{A}^2$ be a system of representants of the different isomorphism classes of cones. Thus, for every $w \in \mathcal{A}^*_{\geq 2}$ there is some unique $ab \in \mathcal{J}$ such that $C([w]) = C(ab)$. Then we write $\tau\big(C(w)\big) := ab$ for the *cone type* (or *isomorphism class*) of the cone $C(w)$. The *boundary* of $C(w)$ is given by the set

$$\partial C(w) = \bigl\{w_0 \in C(w) \bigm| |w_0| = |w|, \exists w' \in \mathcal{A}^* \setminus C(w) : p(w, w') > 0\bigr\}.$$

We have $\{[w] \mid w \in \partial C(w_1)\} = \{[w] \mid w \in \partial C(w_2)\}$ for two ismorphic cones $C(w_1)$ and $C(w_2)$ with $w_1, w_2 \in \mathcal{A}^*_{\geq 2}$, which follows from the following fact: if $x_1 \in \partial C(w_1)$ and $w' \in \mathcal{A}^* \setminus C(w_1)$ with $p(x_1, w') > 0$, then there is, due to 4.1.(1), some $x_2 \in C(w_2)$ with $[x_1] = [x_2]$ and $p([x_2], a) = p([x_1], a) > 0$, where $a \in \mathcal{A}$ is the last letter of w'. This implies existence of some $w'' \in \mathcal{A}^* \setminus C(w_2)$ with $p(x_2, w'') > 0$.

We say that the graph $\mathcal{G}$ is *expanding* if each cone $C(w_0)$, $w_0 \in \mathcal{L}$, contains two proper disjoint subcones, that is, if there exist subcones $C(w_1), C(w_2) \subsetneq C(w_0)$, $w_1, w_2 \in \mathcal{L}$, with $C(w_1) \cap C(w_2) = \emptyset$. We call the random walk *expanding* if the associated graph $\mathcal{G}$ is expanding. The results below do *not* depend on whether the random walk is expanding or not. At the end, however, we will see that the non-expanding case leads to zero entropy.

Finally, let us remark that in the case of K-dependent random walks on $\mathcal{A}^*$ suffix-irreducibility can be defined analogously and cones can be defined in the exactly same way; the different cone types would be defined by words of length K. In Appendix B we will check that suffix-irreducibility and the the "expanding" property are inherited by the blocked letter language if these properties are satisfied for the K-dependent random walk.

4.2. **Covering of Cones by Subcones.** The next task is to cover (up to a finite complement) any cone $C(w)$, $w \in \mathcal{L}$, by a finite set of pairwise disjoint subcones $C_1, \dots, C_{n(w)} \subset C(w)$ such that

$$\bigl\{\tau(C_1), \dots, \tau(C_{n(w)})\bigr\} = \mathcal{J} \quad \text{and} \quad \Bigl|C(w) \setminus \bigcup_{i=1}^{n(w)} C_i\Bigr| < \infty,$$

that is, every cone type appears among these subcones and the subcones cover $C(w)$ up to finitely many words. We then call $C_1, \dots, C_{n(w)}$ a *covering* of $C(w)$. In the next subsection we show how to construct this covering when $\mathcal{G}$ is expanding; in Subsection 4.2.2 we consider the case when $\mathcal{G}$ is *not* expanding.

4.2.1. *Covering for Expanding Random Walks.* Suppose we are given a cone $C(w)$ with $w = w_0 a_0 b_0 \in \mathcal{L}$, where $w_0 \in \mathcal{A}^*$ and $a_0 b_0 \in \mathcal{A}^2$. Inside this cone we can find subcones of the form $C(w_0 w' ab)$ for *each* $ab \in \mathcal{A}^2$ with suitable $w' \in \mathcal{A}^* \setminus \{o\}$ depending on ab due to suffix-irreducibility. Now we want to find subcones of each type $ab \in \mathcal{J}$ which are even pairwise disjoint. We proceed as follows to find these pairwise disjoint cones of all types: since we assume in this subsection that $\mathcal{G}$ is expanding there are paths from $w = w_0 a_0 b_0$ inside $\mathcal{A}^*_{\geq |w|}$ to words $w_0 w_1 a_1 b_1$ and $w_0 w_2 a_2 b_2$, where $w_1, w_2 \in \mathcal{A}^* \setminus \{o\}$, $a_1 b_1, a_2 b_2 \in \mathcal{A}^2$ and $C(w_0 w_1 a_1 b_1) \cap C(w_0 w_2 a_2 b_2) = \emptyset$. Then we have found a subcone of type $\tau(C(a_1 b_1))$, and we search for other cone types in the subcone $C(w_0 w_2 a_2 b_2)$ in the same way. Obviously, a subcone in $C(w_0 w_2 a_2 b_2)$ does not intersect $C(w_0 w_1 a_1 b_1)$. Iterating this step leads to a finite set $\{C_1, \dots, C_{|\mathcal{J}|}\}$ of subcones of $C(w)$ such that $\{\tau(C_1), \dots, \tau(C_{|\mathcal{J}|})\} = \mathcal{J}$ and $C_i \cap C_j = \emptyset$ for $i, j \in \{1, \dots, |\mathcal{J}|\}$ with $i \neq j$. After we have found these non-intersecting subcones of all types in $C(w)$ we cover the cone $C(w)$ by further disjoint subcones: let be

$D = 1 + \max\{|w'| \mid w' \in \bigcup_{i=1}^{|\mathcal{J}|} \partial C_i\}$; define $M_D = \{w' \in C(w) \mid |w'| = D\}$. Then we can choose a subset $M := \{w'_1, \dots, w'_k\} \subseteq M_D$ such that for all $i, j \in \{1, \dots, k\}$ with $i \neq j$ and all $n \in \{1, \dots, |\mathcal{J}|\}$ we have: $C(w'_i) \cap C_n = \emptyset$, $C(w'_i) \cap C(w'_j) = \emptyset$ and

$$C(w) \setminus \left(\bigcup_{m=1}^{|\mathcal{J}|} C_m \cup \bigcup_{n=1}^{k} C(w'_n) \right)$$

is finite. This is done as follows: write $M_D := \{x_1, \dots, x_N\}$ and set $M_0 := \emptyset$. For every $i \in \{1, \dots, N\}$, perform the following steps with increasing i: if $x_i \in \bigcup_{j=1}^{|\mathcal{J}|} C_j \cup \bigcup_{x \in M_{i-1}} C(x)$, then drop x_i and set $M_i := M_{i-1}$. Otherwise, set $M_i := M_{i-1} \cup \{x_i\}$. In the latter case we cannot have $C_j \subset C(x_i)$ for some $j \in \{1, \dots, |\mathcal{J}|\}$ due to the choice of D (words in ∂C_j have word length smaller than D and all words in $C(x_i)$ have length of at least D) and also not $C(x_i) \subset C_j$, which would lead to the contradiction $x_i \in C_j$ otherwise. We also cannot have $C(x_j) \subset C(x_i)$ for $j < i$ because this implies, by Lemma 4.2, $C(x_i) = C(x_j)$ and therefore $x_i \in C(x_j)$. At the end of this procedure we get some M_N and set $M := M_N$. Since every path from w to infinity inside $C(w)$ has to pass through a word of length D we have ensured that each $w' \in C(w)$ with $|w'| = D$ lies in one of the cones $C_1, \dots, C_{\mathcal{J}}, C(x)$, $x \in M$. Thus, the set $C(w) \setminus \bigcup_{m=1}^{|\mathcal{J}|} C_m \cup \bigcup_{x \in M} C(x)$ is finite and the covering of $C(w)$ is given by the subcones

$$C_1, \dots, C_{|\mathcal{J}|}, C(x), x \in M.$$

See Figure 1 for better visualization.

The crucial point now is that we fix a covering for each cone type such that the relative positions of the subcones in the covering of some cone $C(w)$ do *not* depend on the choice of the specific root $w \in \mathcal{L}$ on the boundary of $C(w)$ but only on $\tau(C(w))$: first, for each $ab \in \mathcal{J}$, choose any $w_{ab} \in \mathcal{A}^*$ such that $w_{ab}ab \in \mathcal{L}$ and fix some covering for $C(w_{ab}ab)$, say the cones $C(w_{ab}v_1), \dots, C(w_{ab}v_k)$, where $v_1, \dots, v_k \in \mathcal{A}^*_{\geq 3}$. If $w = w_0 a_1 b_1 \in \mathcal{L}$ with $w_0 \in \mathcal{A}^*$, $a_1 b_1 \in \mathcal{A}^2$ and $\tau(C(w)) = ab = \tau(C(w_{ab}ab))$ then we set the covering of $C(w)$ as the one which is inherited from the covering of $C(w_{ab}ab)$ by the relative location of the subcones, that is, we set the covering of $C(w)$ as the set of subcones $C(w_0 v_1), \dots, C(w_0 v_k)$.

Lemma 4.3. *The set of subcones $C(w_0 v_1), \dots, C(w_0 v_k)$ is a covering of $C(w)$.*

Proof. First, $C(w_0 v_1), \dots, C(w_0 v_k)$ are subcones of $C(w)$ since $ab \in C([w])$ (yielding $w_0 ab \in \partial C(w)$) and due to the following conclusion: for each $i \in \{1, \dots, k\}$, there is a path from $w_{ab}ab$ to $w_{ab}v_i$ through words in $\mathcal{A}^*_{\geq |w_{ab}ab|}$, which implies that there is a path from ab to v_i through words in $\mathcal{A}^*_{\geq 2}$ yielding existence of a path from $w = w_0[w]$ via $w_0 ab$ to $w_0 v_i$ through words in $\mathcal{A}^*_{\geq |w|}$. That is, $C(w_0 v_i) \subset C(w)$.

Since $\mathcal{J} = \{\tau(C(v_1)), \dots, \tau(C(v_k))\}$ the set of subcones $\{C(w_0 v_1), \dots, C(w_0 v_k)\}$ contains all different types. The next step is to show disjointness of the cones $C(w_0 v_1), \dots, C(w_0 v_k)$. Assume w.l.o.g. that $C(w_0 v_1) \subsetneq C(w_0 v_2)$. Then there exists a path from $w_0 v_2$ to $w_0 v_1$ through words in $\mathcal{A}^*_{\geq |w_0 v_2|}$. This implies that there exists a path from v_2 to v_1 through words in $\mathcal{A}^*_{\geq |v_2|} \subseteq \mathcal{A}^*_{\geq 3}$, which implies that there exists a path from $w_{ab}v_2$ to $w_{ab}v_1$ through words in $\mathcal{A}^*_{\geq |w_{ab}v_2|}$ yielding $C(w_{ab}v_1) \subseteq C(w_{ab}v_2)$, a contradiction to the choice of $C(w_{ab}v_1)$,

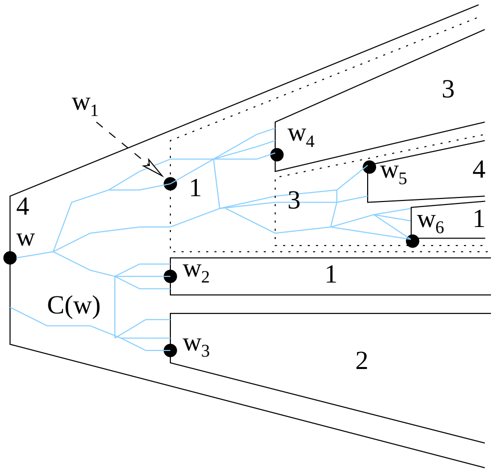


FIGURE 1. Covering of cones by subcones: the numbers represent the four different cone types; the cones with the solid boundary lines belong to the covering of $C(w)$. The construction of a covering is done as follows: e.g., we find three three cones in $C(w)$ whose union covers $C(w)$ up to a finite set, say the cones $C(w_1)$ (type 1), $C(w_2)$ (type 1) and $C(w_3)$ (type 2). We keep the cones $C(w_2)$ and $C(w_3)$ for the covering of $C(w)$ and search for cones of type 3 and 4 in the subcone $C(w_1)$. After having found cones of type 3 and 4 in $C(w_1)$ (for instance, the cones $C(w_4)$ and $C(w_5)$) we take additional disjoint cones in $C(w_1)$ (in the picture the innermost type-1 cone $C(w_6)$ only) into the covering such that the complement of the union of all subcones in the covering is finite. That is, the covering of $C(w)$ consists of the cones $C(w_2)$, $C(w_3)$, $C(w_4)$, $C(w_5)$ and $C(w_6)$.

$C(w_{ab}v_2)$ in the covering of $C(w_{ab}ab)$. Thus, the cones $C(w_0v_1), \ldots, C(w_0v_k)$ are pairwise disjoint.

Analogously, we show that $C(w) \setminus \bigcup_{i=1}^k C(w_0v_i)$ is finite. Assume that this set difference is *not* finite. Then for every $N \in \mathbb{N}$ with $N \geq 3$, there exists some $\bar{w}_N \in \mathcal{A}^*$ with $|\bar{w}_N| = N$ and $w_0\bar{w}_N \in \mathcal{A}^* \cap \overline{\bigcup_{i=1}^k C(w_0v_i)}$ such that there is a path from $w = w_0[w]$ to $w_0\bar{w}_N$ through words in $\mathcal{A}^*_{\geq |w|}$. Since $[w] \in C(ab)$ there is a path from ab to $[w]$ through words in $\mathcal{A}^*_{\geq 2}$ implying that there exists a path from ab to $\bar{w}_N \in \overline{\bigcup_{i=1}^k C(v_i)}$ through words in $\mathcal{A}^*_{\geq 2}$. But this implies that there exists a path from $w_{ab}ab$ to $w_{ab}\bar{w}_N \in \overline{\bigcup_{i=1}^k C(w_{ab}v_i)}$ through words in $\mathcal{A}^*_{\geq |w_{ab}ab|}$. This gives a contradiction since $C(w_{ab}ab) \setminus \bigcup_{i=1}^k C(w_{ab}v_i)$ is finite and therefore N cannot be large. This yields the claim. □

Hence, the covering of a cone depends only on its cone type, which describes the relative location of its subcones in its interior.

We can also cover $\mathcal{L}$ (up to a finite set) by a finite number of non-intersecting subcones, where each cone type appears. To this end, we just apply the algorithm explained above and take pairwise disjoint cones of the form $C(w)$ with $w \in \mathcal{L}$ and $|w| \geq 2$. We denote by $C_1^{(0)}, \dots, C_{n_0}^{(0)}$ the covering of $\mathcal{L}$, which contains all types in $\mathcal{J}$ and which satisfies $\left|\mathcal{L} \setminus \bigcup_{i=1}^{n_0} C_i^{(0)}\right| < \infty$.

4.2.2. *Non-Expanding Random Walks.* Now we explain how to proceed if $\mathcal{G}$ is *not* expanding, that is, there is a cone $C(w)$, $w \in \mathcal{L}$, which does *not* contain two proper disjoint subcones. Recall that due to suffix-irreducibility there is, for every $ab \in \mathcal{J}$, a subcone $C(w_1) \subset C(w)$ with $[w_1] = ab$. Thus, all cones do *not* have two proper disjoint subcones, because otherwise we get a contradiction to the choice of w. This non-expanding case may, in particular, occur if $\mathcal{L}$ is a proper subset of $\mathcal{A}^*$. Take now disjoint cones $C(a_1b_1), \dots, C(a_db_d)$, where $d \in \mathbb{N}$, $a_1b_1, \dots, a_db_d \in \mathcal{A}^2$ with $C(a_ib_i) \cap C(a_jb_j) = \emptyset$ for all $i, j \in \{1, \dots, d\}$ with $i \neq j$ and $\mathcal{L} \setminus \bigcup_{k=1}^d C(a_kb_k)$ is finite. As already mentioned above the cones $C(a_ib_i)$, $i \in \{1, \dots, d\}$, do *not* contain two proper disjoint subcones. Thus, we can then cover any cone $C(w)$, $w \in \mathcal{A}^*_{\geq 2}$, by the subcone $C(w_1)$ for any $w_1 \in C(w)$ with $|w_1| = |w| + 1$ and $p(w, w_1) > 0$.

Example 4.4. *In order to illustrate this situation we give a short example for this case: let* $\mathcal{A} = \{a, b\}$, $p(o,a) = p(a,o) = p(o,b) = p(b,o) = p(a,ab) = p(b,ba) = \frac{1}{2}$ *and* $p(ab,aba) = \frac{2}{3}, p(ba,b) = \frac{1}{3}$, $p(ba,bab) = \frac{3}{4}, p(ab,a) = \frac{1}{4}$. *The set* $\mathcal{L}$ *is then given by all words of the form ababa ... ba, ababa ... bab, baba ... bab and baba ... baba. The random walk is transient and satisfies the Assumptions 2.1 and 2.4. We have* $C(ab) \cap C(ba) = \emptyset$ *and* $C(ab) = C(aba) \cup \{ab\}$ *and* $C(ba) = C(bab) \cup \{ba\}$.

The next step is to show that a non-expanding random walk converges to one of finitely many infinite words. More precisely, since we consider transient random walks, $|X_n|$ tends almost surely to infinity. Therefore, we must have that the prefixes of arbitrary length of X_n stabilize for n large enough, that is, for each $N \in \mathbb{N}$ there exists almost surely some index $n_N \in \mathbb{N}$ such that the prefixes of length N of $X_{n_N}, X_{n_N+1}, X_{n_N+2}, \dots,$ remain constant forever. Thus, $(X_n)_{n\in\mathbb{N}_0}$ tends to some infinite (random) word $X_\infty \in \mathcal{A}^{\mathbb{N}}$.

Lemma 4.5. *If* $(X_n)_{n\in\mathbb{N}_0}$ *is non-expanding, then the support of* X_∞ *is finite.*

Proof. First, assume that X_∞ starts with positive probability with the letter $a_0 \in \mathcal{A}$. Assume also that $\mathbb{P}[\forall n \geq 1 : X_n \in C(a_0b_0c_0) \mid X_0 = a_0b_0c_0] > 0$ for some $b_0c_0 \in \mathcal{A}^2$ with $a_0b_0c_o \in \mathcal{L}$. We denote by A the event that X_∞ starts with the letter a_0 and that the random walk finally enters $C(a_0b_0c_0)$ on its way to infinity. Then $\mathbb{P}[A] > 0$. On this event A, assume now that the random walk tends with positive probability to some infinite words with prefixes wa_1 and wa_2, where $w \in \mathcal{A}^*_{\geq 2}$ starts with the letter a_0 and $a_1, a_2 \in \mathcal{A}$ with $a_1 \neq a_2$. Then there must be words $wa_1b_1c_1, wa_2b_2c_2 \in C(a_0b_0c_0)$, $b_1c_1, b_2c_2 \in \mathcal{A}^2$, such that

$$\mathbb{P}\big[\exists n \in \mathbb{N} : X_n = wa_ib_ic_i, \forall m \geq n : X_m \in C(wa_ib_ic_i) \,\big|\, A\big] > 0 \text{ for } i \in \{1, 2\}.$$

Obviously, $C(wa_1b_1c_1) \cap C(wa_2b_2c_2) = \emptyset$. But this leads to the contradiction that $C(a_0b_0c_0)$ has two proper disjoint subcones. Therefore, $C(wa_1b_1c_1) \cap \mathcal{L} = \emptyset$ or $C(wa_2b_2c_2) \cap \mathcal{L} = \emptyset$, yielding that the letter a_1 (or a_2) is deterministic on the event A. By induction, the infinite limiting word X_∞ is deterministic on the event A, and it depends only on a_0 and b_0c_0. Since there are only finitely many possibilities for a_0 and b_0c_0, the limiting word X_∞ can only take finitely many values. □

The last lemma and suffix-irreducibility directly imply that the support of the random walk is a proper subset of $\mathcal{A}^*$ if $(X_n)_{n\in\mathbb{N}_0}$ is non-expanding. The limiting words in Example 4.4 are $ababab\ldots$ and $bababa\ldots$.

5. Last Entry Times

In this section we prove a law of large numbers, which turns out to describe the asymptotic entropy in the later section. For this purpose, we define last entry times (compare with [7]), for which we derive a law of large numbers. In this section we will assume that $(X_n)_{n\in\mathbb{N}_0}$ is transient and we will assume Assumptions 2.1 and 2.4, where we make explicit comments when these assumptions are essential at some points. Throughout this section, we will also use the following notations: $w_0, w_1, w_2 \in \mathcal{A}^* \setminus \{o\}$ and $a, b, c, d, a_1, b_1, a_2, b_2, \ldots \in \mathcal{A}$.

5.1. **Last Entry Time Process.** We define the following *last entry times*. Let $\mathbf{e}_0$ be the first time at which the random walk visits $\bigcup_{i=1}^{n_0} \partial C_i^{(0)}$ and stays in one of the cones $C_1^{(0)}, \ldots, C_{n_0}^{(0)}$ afterwards forever, that is,

$$\mathbf{e}_0 := \inf\bigl\{m \in \mathbb{N}_0 \,\big|\, \exists i \in \{1, \ldots, n_0\}\, \forall n \geq m : X_n \in C_i^{(0)}\bigr\}.$$

In particular, $X_{\mathbf{e}_0} \in \bigcup_{i=1}^{n(0)} \partial C_i^{(0)}$ and $X_{\mathbf{e}_0 - 1} \notin \bigcup_{i=1}^{n(0)} C_i^{(0)}$. In other words, at time $\mathbf{e}_0$ the random walk finally enters one of the cones $C_i^{(0)}$ with no further exits. Inductively, if $X_{\mathbf{e}_k} = w \in \mathcal{L}$ for $k \geq 0$ and if $C(w)$ has the covering (determined only by the type of $C(w)$) consisting of the subcones $C_1^{(k)}, \ldots, C_{n(w)}^{(k)}$ as explained in Section 4, then

$$\mathbf{e}_{k+1} := \inf\bigl\{m > \mathbf{e}_k \,\big|\, \exists i \in \{1, \ldots, n(w)\}\, \forall n \geq m : X_n \in C_i^{(k)}\bigr\}.$$

In particular, $X_{\mathbf{e}_{k+1}} \in \bigcup_{i=1}^{n(w)} \partial C_i^{(k+1)}$ and $X_{\mathbf{e}_{k+1}-1} \notin \bigcup_{i=1}^{n(w)} \partial C_i^{(k)}$. Transience of $(X_n)_{n\in\mathbb{N}_0}$ yields $\mathbf{e}_k < \infty$ for all $k \in \mathbb{N}_0$ almost surely. Observe that X_n, $n \geq \mathbf{e}_k$, has the prefix w_0 if $X_{\mathbf{e}_k} = w_0ab$. Define the *relative increments* $(\mathbf{W}_k)_{k\in\mathbb{N}_0}$ between two last entry times as follows: set $\mathbf{W}_0 := X_{\mathbf{e}_0}$; for $k \geq 1$: if $X_{\mathbf{e}_{k-1}} = w_0ab$ and $X_{\mathbf{e}_k} = w_0w_1cd$, then set $\mathbf{W}_k := w_1cd$. Since we have only finitely many different cone types and the subcones of the covering of any cone C are nested at uniformly bounded distance (w.r.t. minimal path lengths) to ∂C, the random variables $\mathbf{W}_k$ can take only finitely many different values. Observe that we can reconstruct the values of the $X_{\mathbf{e}_k}$'s from the values of the $\mathbf{W}_k$'s: if $\mathbf{W}_l = w_la_lb_l$ for $l \leq k$ then $X_{\mathbf{e}_k} = w_0w_1\ldots w_ka_kb_k$.

For $w \in \mathcal{L}$, define

$$\mathcal{S}(w) := \bigcup_{i=1}^{n(w)} \partial C_i,$$

where $C_1,\dots,C_{n(w)}$ is the covering of $C(w)$ according to Section 4. Observe that $\mathcal{S}(w_1)=\mathcal{S}(w_2)$ if $C(w_1)=C(w_2)$. Define for $x=a_1\dots a_k\in\mathcal{A}^*$ and $y=a_1\dots a_{k-2}b_{k-1}b_k\dots b_{k+d}\in C(x)$ with $d\geq 1$ and $d=d(x,y):=|y|-|x|$:

$$\mathbb{L}(x,y):=\sum_{n\geq 0}\mathbb{P}\Bigl[X_n=y,X_{n-1}\notin C(y),\forall m\in\{1,\dots,n\}:X_m\in C(x)\Bigm|X_0=x\Bigr].$$

If $d=1$ then $\mathbb{L}(x,y)=\bar{L}(a_{k-1}a_k,b_{k-1}b_kb_{k+1})$. If $d\geq 2$ then $\mathbb{L}(x,y)$ can be rewritten as

$$\sum_{\substack{y_1,\dots,y_{d-1}\in\mathcal{A}^3:\\ y_i[1]=b_{k-2+i}}}\bar{L}(a_{k-1}a_k,y_1)\cdot\prod_{j=1}^{d-2}\bar{L}(y_j[2]y_j[3],y_{j+1})\cdot\bar{L}(y_{d-1}[2]y_{d-1}[3],b_{k+d-2}b_{k+d-1}b_{k+d});\tag{5.1}$$

the last equation follows from the fact that $\mathbb{L}(x,y)$ depends on x only by its last two letters $a_{k-1}a_k$ and by decomposition of the paths from x to y w.r.t the last times when the sets $\mathcal{A}^k$, $\mathcal{A}^{k+1},\dots,\mathcal{A}^{k+d-1}$ are visited on the way from x to y. That is, the l-th factor in (5.1) corresponds to the part of the path from x to y between the last entry of $\mathcal{A}^*_{\geq k+l-1}$ at the word $a_1\dots a_{k-2}b_{k-1}\dots b_{k+l-3}y_{l-1}[2]y_{l-1}[3]$ and the last entry to $\mathcal{A}^*_{\geq k+l}$ at the word $a_1\dots a_{k-2}b_{k-1}\dots b_{k+l-2}y_l[2]y_l[3]$ (with $y_0[2]y_0[3]=a_{k-1}a_k$ and $y_d=b_{k-2}b_{k-1}b_k$). Moreover, $\mathbb{L}(x,y)=\mathbb{L}(a_{k-1}a_k,b_{k-1}b_k\dots b_{k+d})$.

If $x_1\in\mathcal{L}$, $x_2\in\mathcal{S}(x_1)$ and $x_3\in\mathcal{S}(x_2)$ then

$$\mathbb{L}(x_1,x_3)=\sum_{y\in\partial C(x_2)}\mathbb{L}(x_1,y)\cdot\mathbb{L}(y,x_3)$$

by decomposition w.r.t. the last visit of the set $\partial C(x_2)$ since $C(x_3)\subset C(x_2)\subset C(x_1)$. In particular, if $\mathbb{P}[X_{\mathbf{e}_k}=x_1,X_{\mathbf{e}_{k+1}}=x_2,\dots,X_{\mathbf{e}_{k+l}}=x_{l+1}]>0$ for $x_1,\dots,x_{l+1}\in\mathcal{L}$ then we have

$$\begin{aligned}&\mathbb{P}[X_{\mathbf{e}_k}=x_1,X_{\mathbf{e}_{k+1}}=x_2,\dots,X_{\mathbf{e}_{k+l}}=x_{l+1}]\\ =&\sum_{x_0\in\mathcal{L}\setminus C(x_1)}G(o,x_0|1)\cdot p(x_0,x_1)\cdot\mathbb{L}(x_1,x_2)\cdot\ldots\cdot\mathbb{L}(x_l,x_{l+1})\cdot\xi([x_{l+1}])\end{aligned}\tag{5.2}$$

by decomposition on the final entries of the cones $C(x_1),\dots,C(x_{l+1})$. We obtain the following important observation:

Proposition 5.1. *The process* $\bigl(\mathbf{W}_k\bigr)_{k\geq 1}$ *is a Markov chain with transition probabilities*

$$q(x,y):=\begin{cases}\frac{\xi([y])}{\xi([x])}\mathbb{L}(x,y), & \textit{if } y\in\mathcal{S}(x),\\ 0, & \textit{otherwise.}\end{cases}$$

Proof. Let be $w_0,\dots,w_{k+1}\in\mathcal{A}^*\setminus\{o\}$ such that $w_0\in\bigcup_{j=1}^{n_0}\partial C_j^{(0)}$, $w_{i+1}\in S(w_i)$ for all $i\in\{0,\dots,k\}$ and $\mathbb{P}[\mathbf{W}_0=w_0,\dots,\mathbf{W}_{k+1}=w_{k+1}]>0$. For any such sequence $\underline{w}=(w_0,\dots,w_{k+1})$, we set $x_0(\underline{w}):=w_0$ and inductively: if $x_{k-1}(\underline{w})=y_{k-1}a_{k-1}b_{k-1}$ with $y_{k-1}\in\mathcal{A}^*$ and $a_{k-1}b_{k-1}\in\mathcal{A}^2$ then set $x_k(\underline{w}):=y_{k-1}w_k$. That is, if $\mathbf{W}_k=w_k$ then

$X_{\mathbf{e}_k}=x_k(\underline{w})$. Then:

$$\begin{aligned}
&\mathbb{P}\bigl[\mathbf{W}_1=w_1,\dots,\mathbf{W}_k=w_k\bigr] = \sum_{w_0\in\bigcup_{j=1}^{n_0}\partial C_j^{(0)}} \mathbb{P}\bigl[\mathbf{W}_0=w_0,\dots,\mathbf{W}_k=w_k\bigr]\\
=& \sum_{w_0\in\bigcup_{j=1}^{n_0}\partial C_j^{(0)}} \mathbb{P}\bigl[X_{\mathbf{e}_0}=w_0, X_{\mathbf{e}_1}=x_1(\underline{w}),\dots,X_{\mathbf{e}_k}=x_k(\underline{w})\bigr]\\
=& \sum_{w_0\in\bigcup_{j=1}^{n_0}\partial C_j^{(0)}} \sum_{w'\in\mathcal{L}\setminus C(w_0)} G(o,w'|1)\cdot p(w',w_0)\cdot \prod_{i=1}^k \mathbb{L}(x_{i-1}(\underline{w}),x_i(\underline{w}))\cdot \xi([x_k(\underline{w})])\\
=& \sum_{w_0\in\bigcup_{j=1}^{n_0}\partial C_j^{(0)}} \sum_{w'\in\mathcal{L}\setminus C(w_0)} G(o,w'|1)\cdot p(w',w_0)\cdot \prod_{i=1}^k \mathbb{L}(w_{i-1},w_i)\cdot \xi([w_k]).
\end{aligned}$$

The last equation arises from (5.2) by decomposing the paths by the last entries to the sets ∂C_i, where C_i denotes the cone with $X_{\mathbf{e}_i}\in\partial C_i$. Now we obtain:

$$\begin{aligned}
&\mathbb{P}\bigl[\mathbf{W}_{k+1}=w_{k+1}\mid \mathbf{W}_1=w_1,\dots,\mathbf{W}_k=w_k\bigr]\\
=& \frac{\mathbb{P}\bigl[\mathbf{W}_1=w_1,\dots,\mathbf{W}_k=w_k,\mathbf{W}_{k+1}=w_{k+1}\bigr]}{\mathbb{P}\bigl[\mathbf{W}_1=w_1,\dots,\mathbf{W}_k=w_k\bigr]}\\
=& \frac{\sum_{w_0\in\bigcup_{j=1}^{n_0}\partial C_j^{(0)}} \sum_{w'\in\mathcal{L}\setminus C(w_0)} G(o,w'|1)\cdot p(w',w_0)\cdot \prod_{i=1}^{k+1} \mathbb{L}(w_{i-1},w_i)\cdot \xi([w_{k+1}])}{\sum_{w_0\in\bigcup_{j=1}^{n_0}\partial C_j^{(0)}} \sum_{w'\in\mathcal{L}\setminus C(w_0)} G(o,w'|1)\cdot p(w',w_0)\cdot \prod_{i=1}^{k} \mathbb{L}(w_{i-1},w_i)\cdot \xi([w_k])}\\
=& q(x,y).
\end{aligned}$$

□

Define the set

$$\mathcal{W}_0 := \bigl\{w\in\mathcal{A}^* \bigm| \exists w_0\in\mathcal{A}^*, ab\in\mathcal{A}^2 \text{ with } \mathbb{P}[\mathbf{W}_0=w_0ab,\mathbf{W}_1=w]>0\bigr\}\subseteq \mathcal{A}^*_{\geq 3}.$$

The next lemma desribes the support of the random variables $\mathbf{W}_k$; since the proof contains only elementary, tedious calculations, we omit it at this place and hand it in later in Appendix C.

Lemma 5.2. *For all $k\geq 1$,* $\mathrm{supp}(\mathbb{P}[\mathbf{W}_k=\cdot])=\mathcal{W}_0$.

With the last lemma we can show:

Lemma 5.3. *The Markov chain $(\mathbf{W}_k)_{k\in\mathbb{N}}$ is positive recurrent and aperiodic.*

Proof. Since $\mathcal{W}_0$ is finite it suffices to show that the process $(\mathbf{W}_k)_{k\in\mathbb{N}}$ is irreducible and aperiodic. First we show irreducibility. Let be $w_1=w'a_1b_1, w_2\in\mathcal{W}_0$. Then there is some $w_0a_0b_0\in\bigcup_{j=1}^{n_0}\partial C_j^{(0)}$ such that

$$\begin{aligned}
\mathbb{P}[\mathbf{W}_1=w_2] &\geq \mathbb{P}[X_{\mathbf{e}_0}=w_0a_0b_0,\mathbf{W}_1=w_2]\\
&= \sum_{w'\in\mathcal{L}\setminus C(w_0a_0b_0)} G(o,w')p(w',w_0a_0b_0)\mathbb{L}(w_0a_0b_0,w_0w_2)\xi([w_2])>0.
\end{aligned}$$

In particular, $\mathbb{L}(a_0b_0, w_2) = \mathbb{L}(w_0a_0b_0, w_0w_2) > 0$. By construction of coverings, $C(a_1b_1)$ has a subcone of type $\tau(C(a_0b_0))$ in its covering, say the cone $C(\tilde{w})$ with $\tilde{w} \in C(a_1b_1) \cap \mathcal{W}_0$ and $\mathbb{L}(a_1b_1, \tilde{w}) > 0$. Then:

$$\begin{aligned} \mathbb{P}[\mathbf{W}_3 = w_2 \mid \mathbf{W}_1 = w_1] &\geq q(w_1, \tilde{w}) \cdot q(\tilde{w}, w_2) \qquad (5.3)\\ &= \mathbb{L}(a_1b_1, \tilde{w})\mathbb{L}([\tilde{w}], w_2)\frac{\xi([w_2])}{\xi(a_1b_1)} > 0, \end{aligned}$$

which follows from the fact that $\mathbb{L}([\tilde{w}], w_2) > 0$ due to $[\tilde{w}] \in C(a_0b_0)$ and $\mathbb{L}(a_0b_0, w_2) > 0$ (recall the remark before Lemma 5.2). This proves irreducibility and thus positive recurrence of $(\mathbf{W}_k)_{k\in\mathbb{N}}$.

In order to see aperiodicity of the process $(\mathbf{W}_k)_{k\in\mathbb{N}}$ choose in the proof above $w_1 = w_2$, which yields that the period of $(\mathbf{W}_k)_{k\in\mathbb{N}}$ is either 1 or 2. Now let be $w \in \mathcal{W}_0$ and take any $\hat{w} \in \mathcal{W}_0$ with $q(w, \hat{w}) > 0$. Then according to (5.3) we get

$$\mathbb{P}[\mathbf{W}_4 = w, \mathbf{W}_2 = \hat{w} \mid \mathbf{W}_1 = w] = q(w, \hat{w}) \cdot \mathbb{P}[\mathbf{W}_3 = w \mid \mathbf{W}_1 = \hat{w}] > 0,$$

which implies aperiodicity. □

For sake of better identification of the cones, we now switch to a more suitable representation of cones and coverings. We identify the different cone types by numbers $\mathcal{I} := \{1, \dots, r\} \subset \mathbb{N}$. If $C(w)$ is a cone of type $i \in \mathcal{I}$, then the covering of $C(w)$ (according to Subsection 4.2) has $n(i,j)$ subcones of type $j \in \mathcal{I}$. We denote these subcones of type j by $C_{j_{i,k}} = C_{j_{i,k}}(w) \subset C(w)$ with $1 \leq k \leq n(i,j)$ or we just identify them by $j_{i,1}, \dots, j_{i,n(i,j)}$, which correspond to the subcones of type j with different locations inside $C(w)$. In particular, we choose this enumeration of the subcones of type j in a consistent way: if $C(w_{ab}v_m)$ belongs to the covering of $C(ab)$, $i = \tau(C(ab))$, with $C(w_{ab}v_m)$ being the k-th cone of type j in the covering of $C(ab)$ (identified by $j_{i,k}$ w.r.t. ab), then the k-th subcone of type j in the covering of any cone $C(w_0ab)$ is the subcone $C(w_0v_m)$; compare with the construction of the covering of any cone $C(w)$ starting from the covering of the cone $C(w_{ab}ab)$ in Subsection 4.2. That is, by this enumeration of subcones we ensure that the relative position of $C_{j_{i,k}}(w)$ in the interior of $C(w)$ is always the same for any $w \in \mathcal{L}$ with $i = \tau(C(w))$. We will sometimes omit the root w in the notation of the subcones when it will be clear from the context and when only the relative position of a subcone in some given cone will be of importance.

We now track the random walk's way to infinity by looking which of the cones are finally entered successively. For this purpose, define $\mathbf{i}_k := j_{i,l}$ if $\tau(C(X_{\mathbf{e}_{k-1}})) = i$ and $X_{\mathbf{e}_k} \in \partial C_{j_{i,l}}(X_{\mathbf{e}_{k-1}})$. If we set additionally $\mathbf{i}_0 := C(X_{\mathbf{e}_0})$, then the sequence $(\mathbf{i}_k)_{k\in\mathbb{N}_0}$ tracks the random walk's way to infinity.

At this point we recall the relation between $\mathbf{W}_k$ and $X_{\mathbf{e}_k}$: if $X_{\mathbf{e}_0} = \mathbf{W}_0 = w_0a_0b_0$ and $\mathbf{W}_1 = w_1a_1b_1$ then $X_{\mathbf{e}_1} = w_0w_1a_1b_1$; in general, if $X_{\mathbf{e}_{k-1}} = wa_{k-1}b_{k-1}$ and $\mathbf{W}_k = w_ka_kb_k$ then $X_{\mathbf{e}_k} = ww_ka_kb_k$. That is, there is a natural bijection of trajectories of $(\mathbf{W}_k)_{k\in\mathbb{N}_0}$ and $(X_{\mathbf{e}_k})_{k\in\mathbb{N}_0}$. In particular, the values of the $\mathbf{W}_k$'s determine the values of the $\mathbf{i}_k$'s uniquely, since the last two letters of $\mathbf{W}_{k-1}$ describe $\tau(C(X_{\mathbf{e}_{k-1}}))$ and $\mathbf{W}_k$ describes $\tau(C(X_{\mathbf{e}_k}))$ and the corresponding number in the enumeration of subcones. For a better visualization of the values of $\mathbf{i}_k$, see Figure 2.

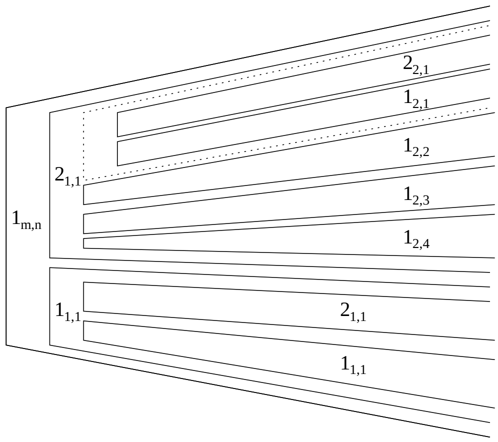


FIGURE 2. Numbering of subcones: the cones with the solid boundary belong to the covering while the cone with the dotted line does not.

In other words, the random variables $\mathbf{i}_k$ collect the information of the different cones which are entered successively by the random walk $(X_n)_{n\in\mathbb{N}_0}$ on its way to infinity, while the $\mathbf{W}_k$'s keep, in addition, the information where the single subcones are finally entered.

Define

$$\mathcal{W} := \left\{ (j_{m,n}, x) \middle| \begin{array}{c} x \in \mathcal{W}_0, \exists w_0 \in \mathcal{L} : \mathbb{P}[\mathbf{W}_0 = w_0, \mathbf{W}_1 = x] > 0, \\ \tau(C([w_0])) = m, \tau(C([x])) = j, 1 \leq n \leq n(m,j) \\ \text{with } x \in \partial C_{j_{m,n}}([w_0]) \end{array} \right\}.$$

In other words, $(j_{m,n}, x) \in \mathcal{W}$ if $x \in \mathcal{W}_0$ with $\tau(C(x)) = j$ and if there is $w_0 a_0 b_0 \in \mathcal{L}$ such that $\tau(C(a_0 b_0)) = m$, $\mathbb{P}[X_{\mathbf{e}_0} = w_0 a_0 b_0, X_{\mathbf{e}_1} = w_0 x] > 0$ and $C(x)$ being the n-th subcone of type j in the covering of $C(a_0 b_0)$.

Proposition 5.4. *The process $\big((\mathbf{i}_k, \mathbf{W}_k)\big)_{k\in\mathbb{N}}$ is a positive recurrent, aperiodic Markov chain on the state space $\mathcal{W}$. Moreover, for $(i_{m,n}, w_1), (j_{s,t}, w_2) \in \mathcal{W}$, the transition probabilities are given by*

$$\mathbb{P}\Big[(\mathbf{i}_k, \mathbf{W}_k) = (j_{s,t}, w_2) \Big| (\mathbf{i}_{k-1}, \mathbf{W}_{k-1}) = (i_{m,n}, w_1)\Big] = \begin{cases} q(w_1, w_2), & \text{if } s = i, \\ 0, & \text{if } s \neq i. \end{cases} \tag{5.4}$$

Proof. Since the values of the $\mathbf{i}_k$'s are uniquely determined by the values of the $\mathbf{W}_k$'s and since the process $(\mathbf{W}_k)_{k\in\mathbb{N}}$ is a Markov chain, we also have that $\big((\mathbf{i}_k, \mathbf{W}_k)\big)_{k\in\mathbb{N}}$ is Markovian with the proposed transition probabilities.

It remains to prove that $\mathrm{supp}(\mathbb{P}[(\mathbf{i}_k, \mathbf{W}_k) = \cdot]) = \mathcal{W}$ for $k \geq 1$ and that $\big((\mathbf{i}_k, \mathbf{W}_k)\big)_{k\in\mathbb{N}}$ is positive recurrent and aperiodic. Since both proofs consist of tedious calculations analogously to the proofs of Lemmas 5.2 and 5.3 we omit these proofs here and refer to Appendix C, where we will hand in them later. □

Let us recall that the values of the $\mathbf{i}_k$'s are uniquely determined by the values of the $\mathbf{W}_k$'s; however, we will explicitely keep the values of the $\mathbf{i}_k$'s in the notation of the process for sake of convenience. Observe that the process $(\mathbf{i}_k)_{k\in\mathbb{N}}$ is, in general, not Markovian. This relies on the fact that $(\mathbf{i}_k)_{k\in\mathbb{N}}$ can be seen as a function of the process $(\mathbf{W}_k)_{k\in\mathbb{N}}$: the values of the $\mathbf{W}_k$'s determine the values of the $\mathbf{i}_k$'s but not vice versa.

Define the following projection for $(i_{k,l}, w_1), (j_{m,n}, w_2) \in \mathcal{W}$:

$$\pi\big((i_{k,l}, w_1), (j_{m,n}, w_2)\big) := \begin{cases} (i, j_{i,n}) =: (i, j_n), & \text{if } m = i, \\ (i, j_{i,1}) = (i, j_1), & \text{if } m \neq i. \end{cases} \tag{5.5}$$

Here, j_l represents the l-th subcone of type j in the covering of a cone of type i, namely the cone represented by $j_{i,l}$. We now define the *hidden Markov chain* $(\mathbf{Y}_k)_{k\in\mathbb{N}}$ by

$$\mathbf{Y}_k := \pi\big((\mathbf{i}_k, \mathbf{W}_k), (\mathbf{i}_{k+1}, \mathbf{W}_{k+1})\big).$$

In other words, $(\mathbf{Y}_k)_{k\in\mathbb{N}}$ traces once again the random walk's way to infinity in terms of which subcones are entered successively *without* distinguishing which of the cone boundary points are the last entry time points $X_{\mathbf{e}_k}$. At this point let us mention that the second branch in the definition of $\pi(\cdot,\cdot)$ is not used for defining $\mathbf{Y}_k$, but it will be of interest in Section 8. Furthermore, observe that $X_{\mathbf{e}_0}$ and $(\mathbf{Y}_k)_{k\in\mathbb{N}}$ allow to reconstruct $(\mathbf{i}_k)_{k\in\mathbb{N}}$.

Define

$$\mathcal{W}_\pi := \big\{(s, t_n) \big| s, t \in \mathcal{I}, 1 \leq n \leq n(s,t)\big\}.$$

That is, t_n corresponds to the n-th subcone of type t in the covering of a cone of type s.

Lemma 5.5. *For all $k \geq 1$, $\mathrm{supp}(\mathbb{P}[\mathbf{Y}_k = \cdot]) = \mathcal{W}_\pi$.*

Proof. The inclusion $\mathrm{supp}(\mathbb{P}[\mathbf{Y}_k = \cdot]) \subset \mathcal{W}_\pi$ is obvious by definition of $\mathbf{Y}_k$ and $\mathcal{W}_\pi$. Now we show the other inclusion. Let be $(s, t_n) \in \mathcal{W}_\pi$. Take any $w_{k-1}a_{k-1}b_{k-1} \in \mathcal{W}_0$ with $\mathbb{P}[\mathbf{W}_{k-1} = w_{k-1}a_{k-1}b_{k-1}] > 0$. Then there exists $w_k a_k b_k \in \mathcal{W}_0$ with $\tau(C(a_k b_k)) = s$ and $q(w_{k-1}a_{k-1}b_{k-1}, w_k a_k b_k) > 0$ due to the construction of coverings. Moreover, there is $w_{k+1}a_{k+1}b_{k+1} \in \mathcal{W}_0$ with $q(w_k a_k b_k, w_{k+1}a_{k+1}b_{k+1}) > 0$ such that $C(w_{k+1}a_{k+1}b_{k+1})$ is the n-th cone of type t in $C(a_k b_k)$. Thus,

$$\begin{aligned} & \mathbb{P}[\mathbf{Y}_k = (s, t_n)] \\ \geq\; & \mathbb{P}[\mathbf{W}_{k-1} = w_{k-1}a_{k-1}b_{k-1}, \mathbf{W}_k = w_k a_k b_k, \mathbf{W}_{k+1} = w_{k+1}a_{k+1}b_{k+1}] \\ =\; & \mathbb{P}[\mathbf{W}_{k-1} = w_{k-1}a_{k-1}b_{k-1}] \cdot q(w_{k-1}a_{k-1}b_{k-1}, w_k a_k b_k) \cdot q(w_k a_k b_k, w_{k+1}a_{k+1}b_{k+1}) > 0, \end{aligned}$$

yielding $(s, t_n) \in \mathrm{supp}(\mathbb{P}[\mathbf{Y}_k = \cdot])$. □

Since the process $(\mathbf{i}_k, \mathbf{W}_k)_{k\in\mathbb{N}}$ is positive recurrent, it has an invariant probability measure ν. Let $(\mathbf{i}_k^{(\nu)}, \mathbf{W}_k^{(\nu)})_{k\in\mathbb{N}}$ be a Markov chain with transition probabilities given by (5.4) but with initial distribution ν. The corresponding hidden Markov chain $(\mathbf{Y}_k^{(\nu)})_{k\in\mathbb{N}}$ is given by

$$\mathbf{Y}_k^{(\nu)} := \pi\big((\mathbf{i}_k^{(\nu)}, \mathbf{W}_k^{(\nu)}), (\mathbf{i}_{k+1}^{(\nu)}, \mathbf{W}_{k+1}^{(\nu)})\big).$$

In the next section we will link the hidden Markov chains $(\mathbf{Y}_k)_{k\in\mathbb{N}}$ and $(\mathbf{Y}_k^{(\nu)})_{k\in\mathbb{N}}$.

5.2. **Entropy of the Hidden Markov Chain related to the Last Entry Time Process.** In this subsection we derive existence of the asymptotic entropy of the hidden Markov chains $(\mathbf{Y}_k^{(\nu)})_{k\in\mathbb{N}}$ and $(\mathbf{Y}_k)_{k\in\mathbb{N}}$.

First, consider the hidden Markov chain $(\mathbf{Y}_k^{(\nu)})_{k\in\mathbb{N}}$: this process is stationary and ergodic since the underlying Markov chain $\big(\mathbf{i}_k^{(\nu)},\mathbf{W}_k^{(\nu)}\big)_{k\in\mathbb{N}}$ is stationary, positive recurrent and aperiodic. Hence, there is a constant $H(\mathbf{Y})\geq 0$ such that

$$\lim_{k\to\infty} -\frac{1}{k}\log \mathbb{P}[\mathbf{Y}_1^{(\nu)}=\underline{y}_1,\dots,\mathbf{Y}_k^{(\nu)}=\underline{y}_k]=H(\mathbf{Y}) \tag{5.6}$$

for almost every realisation $(\underline{y}_1,\underline{y}_2,\dots)\in\mathcal{W}_\pi^{\mathbb{N}}$ of $(\mathbf{Y}_k^{(\nu)})_{k\in\mathbb{N}}$; see e.g. Cover and Thomas [5, Theorem 16.8.1]. The number $H(\mathbf{Y})$ is called the *asymptotic entropy of the (positive recurrent) process* $(\mathbf{Y}_k^{(\nu)})_{k\in\mathbb{N}}$. We now deduce an analogous statement for the process $(\mathbf{Y}_k)_{k\in\mathbb{N}}$.

Proposition 5.6. *For almost every realisation $(\underline{y}_1,\underline{y}_2,\dots)\in\mathcal{W}_\pi^{\mathbb{N}}$ of $(\mathbf{Y}_k)_{k\in\mathbb{N}}$,*

$$\lim_{k\to\infty} -\frac{1}{k}\log \mathbb{P}\big[\mathbf{Y}_1=\underline{y}_1,\dots,\mathbf{Y}_k=\underline{y}_k\big]=H(\mathbf{Y}).$$

Proof. The processes $(\mathbf{Y}_k^{(\nu)})_{k\in\mathbb{N}}$ and $(\mathbf{Y}_k)_{k\in\mathbb{N}}$ differ only by the inital distributions of $(\mathbf{i}_1^{(\nu)},\mathbf{W}_1^{(\nu)})$ and $(\mathbf{i}_1,\mathbf{W}_1)$. Moreover, there are constants $c,C>0$ such that

$$c\cdot\mathbb{P}[(\mathbf{i}_1,\mathbf{W}_1)=(i_{m,n},x)]\leq \nu(i_{m,n},x)\leq C\cdot \mathbb{P}[(\mathbf{i}_1,\mathbf{W}_1)=(i_{m,n},x)]$$

for all $(i_{m,n},x)\in\mathcal{W}$. Denote by μ_1 the distribution of $(\mathbf{i}_1,\mathbf{W}_1)$. We now get for almost every trajectory $(\underline{y}_1,\underline{y}_2,\dots)\in\mathcal{W}_\pi^{\mathbb{N}}$ of $(\mathbf{Y}_k)_{k\in\mathbb{N}}$:

$$\begin{aligned}
H(\mathbf{Y}) &= \lim_{k\to\infty} -\frac{1}{k}\log \mathbb{P}\big[\mathbf{Y}_1^{(\nu)}=\underline{y}_1,\dots,\mathbf{Y}_k^{(\nu)}=\underline{y}_k\big]\\
&= \lim_{k\to\infty} -\frac{1}{k}\log \sum_{\substack{\underline{w}_1,\dots,\underline{w}_{k+1}\in\mathcal{W}:\\ \pi(\underline{w}_j,\underline{w}_{j+1})=\underline{y}_j\\ \text{for } 1\leq j\leq k}} \nu(\underline{w}_1)\mathbb{P}[(\mathbf{i}_l,\mathbf{W}_l)=\underline{w}_l \text{ for } 2\leq l\leq k+1 \mid (\mathbf{i}_1,\mathbf{W}_1)=\underline{w}_1]\\
&= \lim_{k\to\infty} -\frac{1}{k}\log \sum_{\substack{\underline{w}_1,\dots,\underline{w}_{k+1}\in\mathcal{W}:\\ \pi(\underline{w}_j,\underline{w}_{j+1})=\underline{y}_j\\ \text{for } 1\leq j\leq k}} \mu_1(\underline{w}_1)\mathbb{P}[(\mathbf{i}_l,\mathbf{W}_l)=\underline{w}_l \text{ for } 2\leq l\leq k+1 \mid (\mathbf{i}_1,\mathbf{W}_1)=\underline{w}_1]\\
&= \lim_{k\to\infty} -\frac{1}{k}\log \sum_{\substack{\underline{w}_1,\dots,\underline{w}_{k+1}\in\mathcal{W}:\\ \pi(\underline{w}_j,\underline{w}_{j+1})=\underline{y}_j\\ \text{for } 1\leq j\leq k}} \mathbb{P}\big[(\mathbf{i}_1,\mathbf{W}_1)=\underline{w}_1,\dots,(\mathbf{i}_{k+1},\mathbf{W}_{k+1})=\underline{w}_{k+1}\big]\\
&= \lim_{k\to\infty} -\frac{1}{k}\log \mathbb{P}\big[\mathbf{Y}_1=\underline{y}_1,\dots,\mathbf{Y}_k=\underline{y}_k\big].
\end{aligned}$$

□

As a consequence we obtain the next statement:

Corollary 5.7.

$$\lim_{k\to\infty} -\frac{1}{k}\int \log \mathbb{P}\bigl[\mathbf{Y}_1=\underline{y}_1,\dots,\mathbf{Y}_k=\underline{y}_k\bigr]\, d\mathbb{P}(\underline{y}_1,\underline{y}_2,\dots) = H(\mathbf{Y}).$$

Proof. Since $|\mathcal{W}|<\infty$ by definition, there is $\varepsilon_0>0$ such that, for all $\underline{w}_1,\underline{w}_2\in\mathcal{W}$,

$\mathbb{P}[(\mathbf{i}_2,\mathbf{W}_2)=\underline{w}_2 \mid (\mathbf{i}_1,\mathbf{W}_1)=\underline{w}_1]>0$ implies $1\geq \mathbb{P}[(\mathbf{i}_2,\mathbf{W}_2)=\underline{w}_2 \mid (\mathbf{i}_1,\mathbf{W}_1)=\underline{w}_1]\geq\varepsilon_0$.

If $(\underline{y}_1,\dots,\underline{y}_k)\in\mathcal{W}_\pi^k$ with $\mathbb{P}[\mathbf{Y}_1=\underline{y}_1,\dots,\mathbf{Y}_k=\underline{y}_k]>0$ then there are $\underline{w}_1,\dots,\underline{w}_{k+1}\in\mathcal{W}$ with $\pi(\underline{w}_j,\underline{w}_{j+1})=\underline{y}_j$ for $1\leq j\leq k$ and $\mathbb{P}\bigl[(\mathbf{i}_1,\mathbf{W}_1)=\underline{w}_1,\dots,(\mathbf{i}_{k+1},\mathbf{W}_{k+1})=\underline{w}_{k+1}\bigr]>0$. Therefore,

$$\begin{aligned}
0 \;&\leq\; -\frac{1}{k}\log\mathbb{P}\bigl[\mathbf{Y}_1=\underline{y}_1,\dots,\mathbf{Y}_k=\underline{y}_k\bigr]\\
&\leq\; -\frac{1}{k}\log\mathbb{P}\bigl[(\mathbf{i}_1,\mathbf{W}_1)=\underline{w}_1,\dots,(\mathbf{i}_{k+1},\mathbf{W}_{k+1})=\underline{w}_{k+1}\bigr]\\
&\leq\; -\frac{1}{k}\log(c\cdot\varepsilon_0^k) = -\frac{1}{k}\log c-\log\varepsilon_0\leq -\log c-\log\varepsilon_0,
\end{aligned}$$

where $c=\min_{\underline{w}\in\mathcal{W}}\mathbb{P}[(\mathbf{i}_1,\mathbf{W}_1)=\underline{w}]$. Therefore, we may exchange integral and limit, which yields the claim together with Proposition 5.6. □

Let be $w\in\mathcal{L}$ with $|w|\geq 2$. Define

$$\hat{l}(w) := -\log \sum_{w'\in\partial C(w)} L(o,w'|1).$$

We obtain the following law of large numbers:

Proposition 5.8.

$$\lim_{k\to\infty}\frac{\hat{l}(X_{\mathbf{e}_k})}{k} = H(\mathbf{Y}) \quad \textit{almost surely.}$$

Proof. Let be $k\in\mathbb{N}$ and assume for the moment that $\mathbf{W}_l=y_la_lb_l$, where $y_l\in\mathcal{A}^*\setminus\{o\}$ and $a_lb_l\in\mathcal{A}^2$ for $0\leq l\leq k$. That is, $X_{\mathbf{e}_l}=y_0y_1\dots y_la_lb_l$. Furthermore, assume that $\mathbf{Y}_1=(j,t^{(1)})$, where $j=\tau(C(a_1b_1))$, and $\mathbf{Y}_l=(s^{(l)},t^{(l)})$ for $2\leq l\leq k$, where the values of $s^{(2)},\dots,s^{(k-1)}$ and $t^{(1)},\dots,t^{(k-1)}$ are determined by the values of $\mathbf{W}_l=y_la_lb_l$.

One can show that, for almost every realisation $(x_1,\underline{y}_1,\underline{y}_2,\dots)$ of $(X_{\mathbf{e}_1},\mathbf{Y}_1,\mathbf{Y}_2,\dots)$,

$$H(\mathbf{Y}) = \lim_{k\to\infty}-\frac{1}{k}\log\mathbb{P}\bigl[C(X_{\mathbf{e}_1})=C(x_1),\mathbf{Y}_1=\underline{y}_1,\dots,\mathbf{Y}_k=\underline{y}_k\bigr]. \tag{5.7}$$

This follows from the fact that there are only finitely many possibilities for $C(X_{\mathbf{e}_1})$ which do not affect the resulting limit. Since the proof of this equation consists of technical reformulations of the involved probabilities we omit it at this place and give it in Lemma C.1 in Appendix C.

Recall from Equation (3.1) that $G(o,w|1)=G(o,o|1)L(o,w|1)$ for all $w\in\mathcal{L}$ and that $\xi(\cdot)$ can only take finitely many (non-zero) values. We now can conclude as follows:

$$\lim_{k\to\infty}\frac{\hat{l}(X_{\mathbf{e}_k})}{k}=\lim_{k\to\infty}-\frac{1}{k}\log\sum_{w'\in\partial C(y_0y_1\dots y_ka_kb_k)}L(o,w'|1)$$

$$=\lim_{k\to\infty}-\frac{1}{k}\log\sum_{bc\in\mathcal{A}^2:bc\in\partial C(a_kb_k)}L(o,y_0y_1\dots y_kbc|1)$$

$$=\lim_{k\to\infty}-\frac{1}{k}\log\Big[\sum_{w_1\in\partial C(y_0y_1a_1b_1)}\sum_{\substack{w_2,\dots,w_k\in\mathcal{W}_0:\\ w_i\in\partial C(y_ia_ib_i)\\ \text{for all } 2\leq i\leq k}}\sum_{\substack{w'\in\mathcal{L}:\\ w'\notin C(w_1)}}L(o,w'|1)p(w',w_1)\prod_{i=2}^{k}\mathbb{L}([w_{i-1}],w_i)\Big]$$

$$=\lim_{k\to\infty}-\frac{1}{k}\log\Big[\sum_{\substack{w_1\in\partial C(y_0y_1a_1b_1);\\ w_2,\dots,w_k\in\mathcal{W}_0:\\ w_i\in\partial C(y_ia_ib_i)\\ \text{for all } 2\leq i\leq k;\\ w'\in\mathcal{L}\setminus C(w_1)}}G(o,w'|1)p(w',w_1)\xi([w_1])\cdot\prod_{i=2}^{k}\frac{\xi([w_i])}{\xi([w_{i-1}])}\mathbb{L}([w_{i-1}],w_i)\Big]$$

$$=\lim_{k\to\infty}-\frac{1}{k}\log\Big[\sum_{w_1\in\partial C(y_0y_1a_1b_1)}\sum_{\substack{w_2,\dots,w_k\in\mathcal{W}_0:\\ w_i\in\partial C(y_ia_ib_i)\\ \text{for all } 2\leq i\leq k}}\mathbb{P}[X_{\mathbf{e}_1}=w_1]q(y_1[w_1],w_2)\prod_{i=3}^{k}q(w_{i-1},w_i)\Big]$$

$$=\lim_{k\to\infty}-\frac{1}{k}\log\mathbb{P}\left[\begin{array}{c}X_{\mathbf{e}_1}\in C(y_0y_1a_1b_1),\mathbf{Y}_1=(j,t^{(1)}),\\ \mathbf{Y}_2=(s^{(2)},t^{(2)}),\dots,\mathbf{Y}_{k-1}=(s^{(k-1)},t^{(k-1)})\end{array}\right]=H(\mathbf{Y}).$$

The last equation follows from (5.7). We remark that the first coordinate of $\mathbf{Y}_1$ describes only the cone type of $X_{\mathbf{e}_1}$ but there may be several distinct cones of the same type $j\in\mathcal{I}$ with $j=\tau(C(X_{\mathbf{e}_1}))$. □

Recall the definition of $l(w)=-\log L(o,w|1)$ for $w\in\mathcal{L}$.

Corollary 5.9.

$$\lim_{k\to\infty}\frac{l(X_{\mathbf{e}_k})}{k}=H(\mathbf{Y})\quad \textit{almost surely.}$$

Proof. It suffices to compare $\hat{l}(X_{\mathbf{e}_k})$ with $l(X_{\mathbf{e}_k})$. Assume for a moment that $X_{\mathbf{e}_k}=w_k$ with $w_k\in\mathcal{L}$ and that $X_{\mathbf{e}_k}$ is on the boundary of the cone C_k. Then, the probability of walking *inside* C_k from any $w'\in\partial C_k$ to any $w-k\in\partial C_k$ (or vice versa) can be bounded from below by some constant ε_0, because the probabilities depend only on $[w_k],[w']\in\mathcal{A}^2$: that is,

$$\mathbb{P}_{w'}[\exists n\in\mathbb{N}:X_n=w_k,\forall m\leq n:X_n\in C(w')]\geq\varepsilon_0.$$

Therefore,

$$\begin{aligned} L(o, X_{\mathbf{e}_k}|1) &\leq \sum_{w' \in \partial C_k} L(o, w'|1) = \hat{l}(X_{\mathbf{e}_k}), \\ \hat{l}(X_{\mathbf{e}_k}) \cdot \varepsilon_0 &\leq \sum_{w' \in \partial C_k} L(o, w'|1) \cdot \mathbb{P}_{w'}[\exists n \in \mathbb{N} : X_n = w_k, \forall m \leq n : X_n \in C(w')] \\ &\leq |\mathcal{A}^2| \cdot L(o, X_{\mathbf{e}_k}|1). \end{aligned}$$

In the second inequality chain we extended paths from o to w' to paths from o to w_k via w' such that each such path is counted at most $|\mathcal{A}^2|$ times. Taking logarithms, dividing by k and letting k tend to infinity yields the claim. □

Now we come to an important law of large numbers. Denote by ν_0 the invariant probabilty measure of the positive recurrent Markov chain $(\mathbf{W}_k)_{k\in\mathbb{N}}$ and define

$$\lambda := \mathbb{E}[|\mathbf{W}_1^{(\nu)}|] - 2 = \sum_{w \in \mathcal{W}_0} \nu_0(w) \cdot \big(|w| - 2\big). \tag{5.8}$$

Then:

Proposition 5.10.

$$\lim_{k\to\infty} \frac{l(X_n)}{n} = \ell \cdot \lambda^{-1} \cdot H(\mathbf{Y}) \quad \textit{almost surely.}$$

Proof. Define

$$\hat{\mathbf{e}}_k := \inf\big\{m \in \mathbb{N} \big| \forall n \geq m : |X_n| = k\big\}.$$

Observe that $\hat{\mathbf{e}}_k - 1 = \sup\big\{m \in \mathbb{N}\big| |X_m| = k-1\big\}$. Transience yields $\hat{\mathbf{e}}_k < \infty$ almost surely for all $k \in \mathbb{N}$. By [7, Proposition 2.3], $k/(\hat{\mathbf{e}}_k - 1)$ tends to the rate of escape ℓ as $k \to \infty$; hence, $k/\hat{\mathbf{e}}_k \to \ell$ as $k \to \infty$. Define the *maximal last entry times* at time $n \in \mathbb{N}$ as

$$\begin{aligned} \mathbf{k}(n) &:= \max\{k \in \mathbb{N} \mid \hat{\mathbf{e}}_k \leq n\}, \\ \mathbf{t}(n) &:= \max\{k \in \mathbb{N} \mid \mathbf{e}_k \leq n\}. \end{aligned}$$

Obviously, $\mathbf{k}(n) \geq \mathbf{t}(n)$ and each last entry time $\mathbf{e}_k$ corresponds (depending on the concrete realization) to exactly one $\hat{\mathbf{e}}_l$ with $l \geq k$. First, we rewrite

$$\frac{l(X_n)}{n} = \frac{l(X_n) - l(X_{\mathbf{e}_{\mathbf{t}(n)}})}{n} + \frac{l(X_{\mathbf{e}_{\mathbf{t}(n)}})}{\mathbf{t}(n)} \cdot \frac{\mathbf{t}(n)}{\mathbf{k}(n)} \cdot \frac{\mathbf{k}(n)}{\hat{\mathbf{e}}_{\mathbf{k}(n)}} \cdot \frac{\hat{\mathbf{e}}_{\mathbf{k}(n)}}{n}. \tag{5.9}$$

Let ε_1 be the minimal occuring positive single-step transition probability. Define

$$D := \max\left\{ |w_2| - |w_1| \,\middle|\, \begin{array}{c} \exists ab \in \mathcal{A}^2 : C(ab) \text{ has covering } C_1, \ldots, C_{n(ab)}, \\ w_1 \in \partial C(ab), w_2 \in \bigcup_{i=1}^{n(ab)} \partial C_i \end{array} \right\} < \infty.$$

Then we have $\hat{\mathbf{e}}_{\mathbf{k}(n)} \geq \mathbf{e}_{\mathbf{t}(n)} \geq \hat{\mathbf{e}}_{\mathbf{k}(n)-D}$ and $n/\mathbf{e}_{\mathbf{t}(n)} \geq 1$. This implies

$$1 \leq \frac{n}{\mathbf{e}_{\mathbf{t}(n)}} \leq \frac{\hat{\mathbf{e}}_{\mathbf{k}(n)+1}}{\hat{\mathbf{e}}_{\mathbf{k}(n)-D}} = \frac{\hat{\mathbf{e}}_{\mathbf{k}(n)+1}}{\mathbf{k}(n)} \frac{\mathbf{k}(n) - D}{\hat{\mathbf{e}}_{\mathbf{k}(n)-D}} \xrightarrow{n\to\infty} \frac{1}{\ell} \cdot \ell = 1 \quad \text{a.s.,} \tag{5.10}$$

which in turn yields $(n-\mathbf{e}_{\mathbf{t}(n)})/n \to 0$ as $n\to\infty$. Thus, the first quotient on the right hand side of (5.9) tends to zero since

$$\begin{aligned} L(o,X_n|1)\cdot \varepsilon_1^{n-\mathbf{e}_{\mathbf{t}(n)}} &\leq L(o,X_{\mathbf{e}_{\mathbf{t}(n)}}|1) \quad \text{(due to weak symmetry)},\\ L(o,X_{\mathbf{e}_{\mathbf{t}(n)}}|1)\cdot \varepsilon_1^{n-\mathbf{e}_{\mathbf{t}(n)}} &\leq L(o,X_n|1). \end{aligned}$$

Here we used the fact that one can walk from $X_{\mathbf{e}_{\mathbf{t}(n)}}$ to X_n (or vice versa) in $n-\mathbf{e}_{\mathbf{t}(n)}$ steps. By Corollary 5.9, $l(X_{\mathbf{e}_{\mathbf{t}(n)}})/\mathbf{t}(n)$ tends to $H(\mathbf{Y})$. On the other hand side, $\hat{\mathbf{e}}_k/k$ tends almost surely to $1/\ell$ and $\hat{\mathbf{e}}_{\mathbf{k}(n)}/n$ tends to 1 almost surely since $1\leq n/\hat{\mathbf{e}}_{\mathbf{k}(n)}\leq n/\mathbf{e}_{\mathbf{t}(n)}\to 1$ by (5.10). It remains to investigate the limit $\lim_{k\to\infty}\mathbf{k}(n)/\mathbf{t}(n)$. Clearly,

$$\frac{\mathbf{k}(n)}{\mathbf{t}(n)} = \frac{|X_{\hat{\mathbf{e}}_{\mathbf{k}(n)}}|}{\mathbf{t}(n)} = \frac{1}{\mathbf{t}(n)}\Bigl(|X_{\mathbf{e}_1}| + \sum_{i=1}^{\mathbf{t}(n)-1}(|X_{\mathbf{e}_{i+1}}|-|X_{\mathbf{e}_i}|) + (|X_{\hat{\mathbf{e}}_{\mathbf{k}(n)}}|-|X_{\mathbf{e}_{\mathbf{t}(n)}}|)\Bigr).$$

Note that $0\leq |X_{\hat{\mathbf{e}}_{\mathbf{k}(n)}}|-|X_{\mathbf{e}_{\mathbf{t}(n)}}|\leq D$ and $0<|X_{\mathbf{e}_1}|\leq D_1$ almost surely for some suitable constant D_1. Thus, it is sufficient to consider

$$\frac{1}{k}\sum_{i=1}^{k}(|X_{\mathbf{e}_{i+1}}|-|X_{\mathbf{e}_i}|) = \frac{1}{k}\sum_{i=1}^{k}\bigl(|\mathbf{W}_i|-2\bigr).$$

Since $(\mathbf{W}_k)_{k\in\mathbb{N}}$ is positive recurrent, the ergodic theorem yields almost surely

$$\lim_{k\to\infty}\frac{1}{k}\sum_{i=1}^{k}\bigl(|\mathbf{W}_i|-2\bigr) = \sum_{w\in\mathcal{W}_0}\nu_0(w)\bigl(|w|-2\bigr) = \lambda.$$

This finishes the proof and gives the proposed formula. □

6. Existence of Entropy

We now link Proposition 5.10 with the asymptotic entropy of the random walk $(X_n)_{n\in\mathbb{N}_0}$. For this purpose, we follow the reasoning of [8]. First, we need the following lemma:

Lemma 6.1. *There is $R>1$ such that $G(w_1,w_2|R)<\infty$ for all $w_1,w_2\in\mathcal{L}$.*

Proof. A simple adaption of the proof of [16, Proposition 8.2] shows that, for $w_1,w_2\in\mathcal{L}$, $G(w_1,w_2|z)$ has radius of convergence $R(w_1,w_2)>1$. At this point we also need the suffix-irreducibility Assumption 2.4; see Subsection A.1 for a comment on how to weaken this assumption. Since we assume the random walk $(X_n)_{n\in\mathbb{N}_0}$ to be irreducible, the radius of convergence is independent from w_1 and w_2; hence, $G(w_1,w_2|R)<\infty$ for all $w_1,w_2\in\mathcal{L}$ and $R=R(w_1,w_2)$. □

Let us remark that we have also $\bar{L}(ab,cde|R)<\infty$, $\overline{G}(ab,cd|R)<\infty$ and $L(o,a|R)<\infty$ for all $a,b,c,d,e\in\mathcal{A}$, since these generating functions are dominanted by Green functions. In the following let be $\varrho\in[1,R)$.

Lemma 6.2. *There are constants D_1 and $D_2>0$ such that for all $m,n\in\mathbb{N}_0$*

$$p^{(m)}(o,X_n)\leq D_1\cdot D_2^n\cdot\varrho^{-m}.$$

Proof. Denote by $\mathcal{C}_\varrho$ the circle with radius ϱ in the complex plane centered at 0. A straightforward computation together with Fubini's Theorem shows for $m \in \mathbb{N}_0$ and $w \in \mathcal{L}$:

$$\frac{1}{2\pi i} \oint_{\mathcal{C}_\varrho} G(o,w|z)\, z^{-m} \frac{dz}{z} = p^{(m)}(o,w);$$

compare with [8, Lemma 3.4]. Since $G(o,w|z)$ is analytic on $\mathcal{C}_\varrho$, we have $|G(o,w|z)| \leq G(o,w|\varrho)$ for all $|z| = \varrho$. Thus,

$$p^{(m)}(o,w) \leq \frac{1}{2\pi} \cdot \varrho^{-m-1} \cdot G(o,w|\varrho) \cdot 2\pi\varrho = G(o,w|\varrho) \cdot \varrho^{-m}.$$

Set $L := 1 \vee \max\bigl\{\bar{L}\bigl(ab, cde|\varrho\bigr) \mid a,b,c,d,e \in \mathcal{A}\bigr\}$, $C_0 := \varrho \cdot G(o,o|\varrho) \cdot \sum_{a\in\mathcal{A}} L(o,a|\varrho)$ and $C_1 = \max\{\overline{G}(ab,cd|\varrho) \mid ab, cd \in \mathcal{A}^2\}$. Equation (3.5) provides for all $w \in \mathcal{L}$ with $|w| \geq 2$

$$G(o,w|\varrho) = G(o,o|\varrho) \cdot L(o,w|\varrho) \leq C_0 \cdot |\mathcal{A}|^{2(|w|-2)} \cdot L^{|w|-2} \cdot C_1.$$

Set $C_2 := C_0 \vee \max\{G(o,w|\varrho) | w \in \mathcal{L}, |w| \leq 2\}$. Since $|X_n| \leq n$, we obtain the proposed inequality by setting $D_1 := C_1 + C_2$ and $D_2 := |\mathcal{A}|^2 \cdot L$:

$$p^{(m)}(o,X_n) \leq D_1 \cdot |\mathcal{A}|^{2|X_n|} \cdot L^{|X_n|} \cdot \varrho^{-m} \leq D_1 \cdot |\mathcal{A}|^{2n} \cdot L^n \cdot \varrho^{-m} = D_1 \cdot D_2^n \cdot \varrho^{-m}.$$

□

The following technical lemma will be used in the proof of the next theorem:

Lemma 6.3. *Let $(A_n)_{n\in\mathbb{N}}$, $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$ be sequences of strictly positive numbers with $A_n = a_n + b_n$. Assume that $\lim_{n\to\infty} -\frac{1}{n}\log A_n = c \in [0,\infty)$ and that $\lim_{n\to\infty} b_n/q^n = 0$ for all $q \in (0,1)$. Then $\lim_{n\to\infty} -\frac{1}{n}\log a_n = c$.*

Proof. A proof can be found in [8, Lemma 3.5]. □

Lemma 6.4. *For $n \in \mathbb{N}$, consider the function $f_n : \mathcal{L} \to \mathbb{R}$ defined by*

$$f_n(w) := \begin{cases} -\frac{1}{n}\log \sum_{m=0}^{n^2} p^{(m)}(o,w), & \text{if } p^{(n)}(o,w) > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Then there are constants d and D such that $d \leq f_n(w) \leq D$ for all $n \in \mathbb{N}$ and $w \in \mathcal{L}$.

Proof. Let be $w \in \mathcal{L}$ and $n \in \mathbb{N}$ with $p^{(n)}(o,w) > 0$. For $w_1 \in \mathcal{L}$ and $z > 0$, define the first return generating function as

$$U(w_1,w_1|z) := \sum_{n\geq 1} \mathbb{P}\bigl[X_n = w_1, \forall m \in \{1,\dots,n-1\} : X_m \neq w_1 \bigl| X_0 = w_1\bigr] \cdot z^n.$$

Recall the number $R > 1$ from Lemma 6.1. Then

$$G(w,w|1) \leq \frac{1}{1-\frac{1}{R}}; \tag{6.1}$$

indeed, since $G(w,w|z) = \bigl(1 - U(w,w|z)\bigr)^{-1}$ it must be that $U(w,w|z) < 1$ for all $w \in \mathcal{L}$ and all $z \in [0,R)$; moreover, $U(w,w|0) = 0$, $U(w,w|z)$ is continuous, strictly increasing and strictly convex for $z \in [0,R)$, so we must have $U(w,w|z) \leq 1/R$ for all $z \in [0,R)$, providing (6.1).

Define $F(o,w) := \sum_{n\geq 0} f^{(k)}(o,w)$, where $f^{(k)}(o,w)$ is the probability of starting at o and with the first visit to w at time k. By conditioning on the first visit to w we get $G(o,w|1) = F(o,w)G(w,w|1)$. Therefore,

$$\sum_{m=0}^{n^2} p^{(m)}(o,w) \leq G(o,w|1) = F(o,w)\cdot G(w,w|1) \leq \frac{1}{1-\frac{1}{R}},$$

that is,

$$f_n(w) \geq -\frac{1}{n}\log\frac{1}{1-\frac{1}{R}} \geq -\log\frac{1}{1-\frac{1}{R}} =: d.$$

For the upper bound, observe that $w\in\mathcal{L}$ with $p^{(n)}(o,w)>0$ can be reached from o in n steps with a probability of at least ε_0^n, where

$$\varepsilon_0 := \min\{p(w_1,w_2) \mid w_1,w_2\in\mathcal{A}^*, p(w_1,w_2)>0\} > 0$$

is independent from w. Thus, the sum $\sum_{m=0}^{n^2} p^{(m)}(o,w)$ has a value greater or equal to ε_0^n. Hence, $f_n(x) \leq -\log\varepsilon_0 =: D$. □

Now we can finally prove:

Proof of Theorem 2.5. Recall Equation (3.1). We can rewrite $\ell\cdot\lambda^{-1}\cdot H(\mathbf{Y})$ as

$$\frac{\ell\cdot H(\mathbf{Y})}{\lambda} = \int \frac{\ell\cdot H(\mathbf{Y})}{\lambda}\,d\mathbb{P} = \int \lim_{n\to\infty} -\frac{1}{n}\log L\big(o,X_n(\omega)\big|1\big)\,d\mathbb{P}(\omega)$$

$$= \int \lim_{n\to\infty} -\frac{1}{n}\log\frac{G\big(o,X_n(\omega)\big|1\big)}{G(o,o|1)}\,d\mathbb{P}(\omega) = \int \lim_{n\to\infty} -\frac{1}{n}\log G\big(o,X_n(\omega)|1\big)\,d\mathbb{P}(\omega).$$

Recall that π_n denotes the distribution of X_n. Since

$$G(o,X_n|1) = \sum_{m\geq 0} p^{(m)}(o,X_n) \geq p^{(n)}(o,X_n) = \pi_n(X_n),$$

we have

$$\frac{\ell\cdot H(\mathbf{Y})}{\lambda} \leq \int \liminf_{n\to\infty} -\frac{1}{n}\log\pi_n\big(X_n(\omega)\big)\,d\mathbb{P}(\omega). \tag{6.2}$$

The next aim is to prove that $\limsup_{n\to\infty} -\frac{1}{n}\mathbb{E}\big[\log\pi_n(X_n)\big] \leq \ell\cdot H(\mathbf{Y})/\lambda$. We now apply Lemma 6.3 by setting

$$A_n := \sum_{m\geq 0} p^{(m)}(o,X_n),\ a_n := \sum_{m=0}^{n^2} p^{(m)}(o,X_n) \text{ and } b_n := \sum_{m\geq n^2+1} p^{(m)}(o,X_n).$$

By Lemma 6.2,

$$b_n \leq \sum_{m\geq n^2+1} D_1\cdot D_2^n\cdot\varrho^{-m} = D_1\cdot D_2^n\cdot\frac{\varrho^{-n^2-1}}{1-\varrho^{-1}}.$$

Therefore, b_n decays faster than any geometric sequence. Applying Lemma 6.3 together with (3.1) gives almost surely

$$\frac{\ell\cdot H(\mathbf{Y})}{\lambda} = \lim_{n\to\infty} -\frac{1}{n}\log L(o,X_n) = \lim_{n\to\infty} -\frac{1}{n}\log G(o,X_n) = \lim_{n\to\infty} -\frac{1}{n}\log\sum_{m=0}^{n^2} p^{(m)}\big(o,X_n\big).$$

Due to Lemma 6.4 we can apply the Dominated Convergence Theorem and get:

$$\frac{\ell \cdot H(\mathbf{Y})}{\lambda} = \int \frac{\ell \cdot H(\mathbf{Y})}{\lambda}\, d\mathbb{P} = \int \lim_{n\to\infty} -\frac{1}{n} \log \sum_{m=0}^{n^2} p^{(m)}(o, X_n)\, d\mathbb{P}$$

$$= \lim_{n\to\infty} \int -\frac{1}{n} \log \sum_{m=0}^{n^2} p^{(m)}(o, X_n)\, d\mathbb{P} = \lim_{n\to\infty} -\frac{1}{n} \sum_{w\in\mathcal{L}} p^{(n)}(o,w) \log \sum_{m=0}^{n^2} p^{(m)}(o,w).$$

For $w \in \mathcal{L}$, define the following distribution μ_0 on $\mathcal{L}$:

$$\mu_0(w) := \frac{1}{n^2+1} \sum_{m=0}^{n^2} p^{(m)}(o,w).$$

Recall that the non-negativity of the Kullback-Leibler divergence (in this context also called *Shannon's Inequality*) gives

$$-\sum_{w\in\mathcal{L}} p^{(n)}(o,w) \log \mu_0(w) \geq -\sum_{w\in\mathcal{L}} p^{(n)}(o,w) \log p^{(n)}(o,w).$$

Therefore,

$$\begin{aligned}\frac{\ell \cdot H(\mathbf{Y})}{\lambda} &\geq \limsup_{n\to\infty} -\frac{1}{n} \sum_{w\in\mathcal{L}} p^{(n)}(o,w) \log(n^2+1) - \frac{1}{n} \sum_{w\in\mathcal{L}} p^{(n)}(o,w) \log p^{(n)}(o,w) \\ &= \limsup_{n\to\infty} -\frac{1}{n} \int \log \pi_n(X_n)\, d\mathbb{P}.\end{aligned}$$

Now we can conclude with (6.2) and Fatou's Lemma:

$$\begin{aligned}\frac{\ell \cdot H(\mathbf{Y})}{\lambda} &\leq \int \liminf_{n\to\infty} -\frac{1}{n} \log \pi_n(X_n) d\mathbb{P} \leq \liminf_{n\to\infty} \int -\frac{1}{n} \log \pi_n(X_n) d\mathbb{P} \\ &\leq \limsup_{n\to\infty} \int -\frac{1}{n} \log \pi_n(X_n) d\mathbb{P} \leq \frac{\ell \cdot H(\mathbf{Y})}{\lambda}.\end{aligned}$$

Thus, the asymptotic entropy $h := \lim_{n\to\infty} -\frac{1}{n}\mathbb{E}\big[\log \pi_n(X_n)\big]$ exists and equals $\ell \cdot H(\mathbf{Y})/\lambda$. □

Finally, we can prove:

Proof of Corollary 2.7. The proofs of the statements in Corollary 2.7 are completely analogous to the proofs in [8, Corollary 3.9, Lemma 3.10], where [8, Lemma 3.10] holds also in the case $h = 0$. □

Proof of Corollary 2.8. Recall the definition of $F(o,w)$ from the proof of Lemma 6.4 and the equation $G(o,w|1) = F(o,w)G(w,w|1)$. This yields together with (3.1):

$$\mathbb{P}[\exists n \in \mathbb{N}_0 : X_n = w] = F(o,w) = \frac{G(o,w|1)}{G(w,w|1)} = \frac{G(o,o|1)}{G(w,w|1)} L(o,w|1).$$

Since $1 \leq G(X_n, X_n|1) \leq 1/(1-\frac{1}{R})$ with R from Lemma 6.1, we obtain the proposed result due to Proposition 5.10. □

7. Calculation of the Entropy

In this section we collect several results about the asymptotic entropy. We show how the entropy can be calculated numerically or even exactly in some special cases, and we give some inequalities.

7.1. **Numerical Calculation and Inequalities.** In order to compute $h = \ell \cdot H(\mathbf{Y})/\lambda$ we have to calculate the three factors: while there are formulas for ℓ (see [7, Theorem 2.4]) and λ (given by (5.8)), it remains to explain how to calculate $H(\mathbf{Y})$. For this purpose, define for random variables $A_1, \dots, A_n$ on a finite state space $\mathcal{W}_A$ the *joint entropy* as

$$H(A_1, \dots, A_n) := -\sum_{a_1,\dots,a_n \in \mathcal{W}_A} \mathbb{P}\bigl[A_1 = a_1, \dots, A_n = a_n\bigr] \log \mathbb{P}\bigl[A_1 = a_1, \dots, A_n = a_n\bigr],$$

and let the *conditional entropy* $H(A_n | A_1, \dots, A_{n-1})$ be defined as

$$-\sum_{a_1,\dots,a_n \in \mathcal{W}_A} \mathbb{P}\bigl[A_1 = a_1, \dots, A_n = a_n\bigr] \log \mathbb{P}\bigl[A_n = a_n \big| A_1 = a_1, \dots, A_{n-1} = a_{n-1}\bigr].$$

Here, we set $0 \cdot \log 0 := 0$, since $x \log x \to 0$ as $x \to 0+$. By Cover and Thomas [5, Theorem 4.2.1], we have $H(\mathbf{Y}) = \lim_{n\to\infty} \frac{1}{n} H(\mathbf{Y}_1^{(\nu)}, \dots, \mathbf{Y}_n^{(\nu)})$. In general, the computation of $H(\mathbf{Y})$ is a hard task. But there is a simple way for a numerical calculation of $H(\mathbf{Y})$, which follows from the inequalities

$$H\bigl(\mathbf{Y}_n^{(\nu)} \big| \bigl((\mathbf{i}_1^{(\nu)}, \mathbf{W}_1^{(\nu)}), (\mathbf{i}_2^{(\nu)}, \mathbf{W}_2^{(\nu)})\bigr), \mathbf{Y}_1^{(\nu)}, \dots, \mathbf{Y}_{n-1}^{(\nu)}\bigr) \leq H(\mathbf{Y}) \leq H(\mathbf{Y}_n^{(\nu)} \mid \mathbf{Y}_1^{(\nu)}, \dots, \mathbf{Y}_{n-1}^{(\nu)}) \tag{7.1}$$

for all $n \in \mathbb{N}$; see [5, Theorem 4.5.1]. In particular, it is even shown that

$$H(\mathbf{Y}_n^{(\nu)} \mid \mathbf{Y}_1^{(\nu)}, \dots, \mathbf{Y}_{n-1}^{(\nu)}) - H\bigl(\mathbf{Y}_n^{(\nu)} \big| \bigl((\mathbf{i}_1^{(\nu)}, \mathbf{W}_1^{(\nu)}), (\mathbf{i}_2^{(\nu)}, \mathbf{W}_2^{(\nu)})\bigr), \mathbf{Y}_1^{(\nu)}, \dots, \mathbf{Y}_{n-1}^{(\nu)}\bigr) \xrightarrow{n\to\infty} 0.$$

Hence, one can calculate $H(\mathbf{Y})$ numerically up to an arbitrarily small error. Obviously, this numerical approach depends strongly on the ability to solve the system of equations given by (3.2).

We now investigate whether the entropy is non-zero or not.

Corollary 7.1. *If the random walk is expanding, then $h > 0$. Otherwise, $h = 0$.*

Proof. Take any $(i_{k,l}, w_1), (j_{p,q}, w_2) \in \mathcal{W}$ with

$$\mathbb{P}[(\mathbf{i}_1^{(\nu)}, \mathbf{W}_1^{(\nu)}) = (i_{k,l}, w_1), (\mathbf{i}_2^{(\nu)}, \mathbf{W}_2^{(\nu)}) = (j_{p,q}, w_2)] > 0.$$

The values $(i_{k,l}, w_1), (j_{p,q}, w_2)$ determine the value of $\mathbf{Y}_1^{(\nu)}$ uniquely. In the expanding case, there are at least two elements $(s_{j,m}, w'), (t_{j,n}, w'') \in \mathcal{W}$ such that $w', w'' \in C([w_2])$ with $C(w') \cap C(w'') = \emptyset$ and $q(w_2, w') > 0$ and $q(w_2, w'') > 0$, yielding $\pi\bigl((j_{p,q}, w_2), (s_{j,m}, w')\bigr) \neq \pi\bigl((j_{p,q}, w_2), (t_{j,n}, w'')\bigr)$. Let w' be in the m-th cone of type s in the covering of $C([w_2])$. Then set

$$\begin{aligned} & P\bigl((i_{k,l}, w_1), (j_{p,q}, w_2), (s_{j,m}, w')\bigr) \\ := \; & \mathbb{P}\bigl[\mathbf{Y}_2^{(\nu)} = \bigl(\tau(C(w_2)), s_m\bigr) \bigm| (\mathbf{i}_1^{(\nu)}, \mathbf{W}_1^{(\nu)}) = (i_{k,l}, w_1), (\mathbf{i}_2^{(\nu)}, \mathbf{W}_2^{(\nu)}) = (j_{p,q}, w_2)\bigr] \\ \geq \; & q(w_2, w') > 0. \end{aligned}$$

Since $q(w_2,w'')>0$ and $C(w')\cap C(w'')=\emptyset$, we also have $P\big((i_{k,l},w_1),(j_{p,q},w_2),(t_{j,n},w'')\big)>0$ implying $P\big((i_{k,l},w_1),(j_{p,q},w_2),(s_{j,m},w')\big)<1$. From (7.1) follows then

$$\begin{aligned} H(\mathbf{Y}) &\geq H\big(\mathbf{Y}_2^{(\nu)}\big|\big((\mathbf{i}_1^{(\nu)},\mathbf{W}_1^{(\nu)}),(\mathbf{i}_2^{(\nu)},\mathbf{W}_2^{(\nu)})\big),\mathbf{Y}_1^{(\nu)}\big)\\ &\geq P\big((i_{k,l},w_1),(j_{p,q},w_2),(s_{j,m},w')\big)\log P\big((i_{k,l},w_1),(j_{p,q},w_2),(s_{j,m},w')\big)>0. \end{aligned}$$

Thus, we have shown that $h>0$ if $(X_n)_{n\in\mathbb{N}_0}$ is expanding.

Now consider the case when the random walk on $\mathcal{L}$ is *not* expanding. Then each cone has a covering consisting of only one single subcone. This implies that the value $\tau(C(\mathbf{W}_1^{(\nu)}))=\mathbf{i}_1^{(\nu)}$ determines uniquely the values $\tau(C(\mathbf{W}_k^{(\nu)}))$ for $k\geq 2$. Moreover, given the value of $\tau(C(\mathbf{W}_1^{(\nu)}))$ the values of $\mathbf{Y}_k^{(\nu)}$, $k\geq 1$, are deterministic. That is, $\mathbf{Y}_n^{(\nu)}$ is uniquely determined by $\mathbf{Y}_1^{(\nu)}$, hence $\mathbb{P}[\mathbf{Y}_n^{(\nu)}=\cdot\mid \mathbf{Y}_1^{(\nu)}=(s,t_n)]\in\{0,1\}$. This implies

$$0\leq H(\mathbf{Y})\leq H(\mathbf{Y}_n^{(\nu)}\mid \mathbf{Y}_1^{(\nu)},\dots,\mathbf{Y}_{n-1}^{(\nu)})\leq H(\mathbf{Y}_n^{(\nu)}\mid \mathbf{Y}_1^{(\nu)})=0,$$

where the last inequality follows from [5, Theorem 2.6.5]. Thus, $h=0$. □

In order to get a complete picture, we show that the entropy is zero for recurrent random walks:

Corollary 7.2. *If $(X_n)_{n\in\mathbb{N}_0}$ is recurrent then $h=0$.*

Proof. Clearly, $-\frac{1}{n}\mathbb{E}\big[\log\pi_n(X_n)\big]\geq 0$. Assume now that $\limsup_{n\to\infty}-\frac{1}{n}\mathbb{E}\big[\log\pi_n(X_n)\big]=c>0$. Then there is a (deterministic) sequence $(n_k)_{k\in\mathbb{N}}$ such that, for any $\varepsilon_1\in(0,c)$,

$$-\frac{1}{n_k}\mathbb{E}\big[\log\pi_{n_k}(X_{n_k})\big]\geq c-\varepsilon_1>0 \tag{7.2}$$

for all $k\in\mathbb{N}$. Denote by ε_0 the minimal occuring positive single-step transition probability of $(X_n)_{n\in\mathbb{N}_0}$. Then $-\frac{1}{n_k}\log\pi_{n_k}(X_{n_k})\leq-\log\varepsilon_0$. Choose $N\in\mathbb{N}$ with $1/N<c-\varepsilon_1$. Then there is some $\delta>0$ with

$$\mathbb{P}\Big[-\frac{1}{n_k}\log\pi_{n_k}(X_{n_k})\geq\frac{1}{N}\Big]\geq\delta\quad\forall k\in\mathbb{N}.$$

To see this, assume that $\delta=\delta_k$ depends on k with $\liminf_{k\to\infty}\delta_k=0$: then we get with (7.2)

$$(-\log\varepsilon_0)\cdot\delta_k+(1-\delta_k)\frac{1}{N}\geq-\frac{1}{n_k}\mathbb{E}\big[\log\pi_{n_k}(X_{n_k})\big]\geq c-\varepsilon_1;$$

If δ_k tends to zero then we get a contradiction to the choice of N.

Choose now $\varepsilon>0$ arbitrarily small with $\varepsilon<\delta$. Since $\ell=0$ in the recurrent case, there is some index $K\in\mathbb{N}$ such that for all $k\geq K$:

$$\delta-\varepsilon\leq\mathbb{P}\big[-\log\pi_{n_k}(X_{n_k})\geq n_k/N,|X_{n_k}|\leq\varepsilon n_k\big]\leq e^{-n_k/N}\cdot|\mathcal{A}|^{\varepsilon n_k}$$

which yields the inequality

$$\frac{1}{N}+\frac{1}{n_k}\log(\delta-\varepsilon)\leq\varepsilon\log|\mathcal{A}|.$$

But this gives a contradiction if we make ε sufficiently small since the right hand side tends to zero, but the left hand side to $\frac{1}{N}$ as $k\to\infty$. Thus, $\limsup_{n\to\infty} -\frac{1}{n}\mathbb{E}\big[\log \pi_n(X_n)\big]=0$, yielding $h=0$. □

Finally, we state an inequality which connects entropy, drift and growth. For this purpose, define $\mathcal{A}^*_{\leq n}=\{w\in\mathcal{A}^*\mid |w|\leq n\}$ for $n>0$. The *growth* of $\mathcal{A}^*$ is then given by $g:=\lim_{n\to\infty}\frac{1}{n}\log|\mathcal{A}^*_{\leq n}|$. Since $|\mathcal{A}^n|\leq|\mathcal{A}^*_{\leq n}|\leq n|\mathcal{A}^n|$, we have $g=\log|\mathcal{A}|$. We get the following connection between entropy, drift and growth:

Theorem 7.3. $h\leq \ell\cdot\log|\mathcal{A}|$.

Proof. Let be $\varepsilon>0$. By Corollary 2.7 (1), there is some $N_\varepsilon\in\mathbb{N}$ such that for all $n\geq N_\varepsilon$:
$$1-\varepsilon\leq \mathbb{P}\big[-\log\pi_n(X_n)\geq (h-\varepsilon)n, |X_n|\leq (\ell+\varepsilon)n\big]\leq e^{-(h-\varepsilon)n}\cdot|\mathcal{A}^*_{\leq(\ell+\varepsilon)n}|.$$
Taking logarithms and dividing by n gives
$$(h-\varepsilon)+\frac{1}{n}\log(1-\varepsilon)\leq(\ell+\varepsilon)\cdot\frac{1}{(\ell+\varepsilon)n}\log|\mathcal{A}^*_{\leq(\ell+\varepsilon)n}|.$$
Making ε arbirtraily small and sending $n\to\infty$ yields the proposed claim. □

Let us remark that similar inequalities have been proved by Kaimanovich and Woess [14] for time and space homogeneous random walks and in [8] for random walks on free products.

7.2. **Exact Formula for Unambiguous Cone Boundaries.** In this subsection we give an exact formula for the asymptotic entropy in some special case. We call $ab\in\mathcal{A}^2$ *unambiguous* if $\partial C(ab)=\{ab\}$. In other words, whenever the random walk enters a subcone of type $C(wab)$, $w\in\mathcal{A}^*$, it must enter it through its single boundary point wab. We call the cone type $\tau(C(ab))$ also unambiguous. Existence of an unambiguous cone allows us to "cut" the random walk into i.i.d. pieces and to obtain a formula for the entropy $H(\mathbf{Y})$. For $n\in\mathbb{N}$, $x_2,\dots,x_n\in\mathcal{W}_0$ and unambiguous $ab\in\mathcal{A}^2$ define
$$\begin{aligned} w(ab,x_2,\dots,x_n) &:= \mathbb{P}\big[\mathbf{W}_2=x_2,\dots,\mathbf{W}_n=x_n,[\mathbf{W}_n]=ab\big|[\mathbf{W}_1]=ab\big],\\ \tilde{w}(ab,x_2,\dots,x_n) &:= \sum_{\substack{y_2,\dots,y_n\in\mathcal{W}_0:\\ y_i\in\partial C(x_i)\\ \text{for } 2\leq i\leq n}} \mathbb{P}\big[\mathbf{W}_2=y_2,\dots,\mathbf{W}_n=y_n,[\mathbf{W}_n]=ab\big|[\mathbf{W}_1]=ab\big],\end{aligned}$$
In particular, $\tilde{w}(ab,x_2)=\mathbb{P}\big[\mathbf{W}_2=x_2,[\mathbf{W}_2]=ab\big|[\mathbf{W}_1]=ab\big]$. Recall that ν denotes the invariant probability measure of the process $(\mathbf{i}_k,\mathbf{W}_k)_{k\in\mathbb{N}}$. For unambiguous $ab\in\mathcal{A}^2$, set
$$\nu_{ab}:=\sum_{(i_{m,n},x)\in\mathcal{W}:[x]=ab}\nu(i_{m,n},x).$$
Then:

Theorem 7.4. *If $ab\in\mathcal{A}^2$ is unambiguous, then*
$$H(\mathbf{Y})=-\nu_{ab}\sum_{n\geq 1}\sum_{\substack{x_2,\dots x_{n-1}\in\mathcal{W}_0:\\ [x_i]\neq ab \text{ for } 2\leq i\leq n-1}}\sum_{\substack{x_n\in\mathcal{W}_0:\\ [x_n]=ab}} w(ab,x_2,\dots,x_n)\log\tilde{w}(ab,x_2,\dots,x_n).$$

Proof. Write $\alpha := \tau(C(ab))$. By Proposition 5.6, we have

$$-\frac{1}{n}\log \mathbb{P}[\mathbf{Y}_1 = \underline{y}_1, \dots, \mathbf{Y}_n = \underline{y}_n] \xrightarrow{n\to\infty} H(\mathbf{Y})$$

for almost every trajectory $(\underline{y}_1, \underline{y}_2, \dots) \in \mathcal{W}_\pi^{\mathbb{N}}$. For any such trajectory, we define

$$N_0 := \min\bigl\{m \in \mathbb{N} \bigm| \tau(\mathbf{W}_{m+1}) = \alpha\bigr\} \text{ and } N_k := \min\bigl\{m \in \mathbb{N} \bigm| m > N_{k-1}, \tau(\mathbf{W}_{m+1}) = \alpha\bigr\}.$$

Define $d(n) := \max\{k \in \mathbb{N}_0 \mid N_k \leq n\}$. Since $\mathbf{Y}_{N_j}$ has the form $(t, \alpha_{t^{(n)},m})$ for some cone type $t \in \mathcal{I}$, $1 \leq m \leq n(t,\alpha)$, and $[\mathbf{W}_{N_k+1}] = ab$ for all $k \in \mathbb{N}$ we can use the strong Markov property as follows for all $n \geq 1$ and almost every trajectory $(\underline{y}_1, \underline{y}_2, \dots) \in \mathcal{W}_\pi^{\mathbb{N}}$:

$$\begin{aligned} & \mathbb{P}\bigl[\mathbf{Y}_{N_j+1} = \underline{y}_{N_j+1}, \dots, \mathbf{Y}_{N_j+n} = \underline{y}_n \mid \mathbf{Y}_1 = \underline{y}_1, \dots, \mathbf{Y}_{N_j} = \underline{y}_{N_j}\bigr] \\ = \; & \mathbb{P}\bigl[\mathbf{Y}_{N_j+1} = \underline{y}_{N_j+1}, \dots, \mathbf{Y}_{N_j+n} = \underline{y}_n \mid [\mathbf{W}_{N_j+1}] = ab\bigr]. \end{aligned}$$

In other words, the $\mathbf{Y}_k$'s collect only the information which cones are entered successively, but we know that the (N_j+1)-th cone is entered through a boundary point with last two letters ab; hence, one can restart the process at some word ending with ab in the above equation without changing probabilities. Therefore, we can rewrite the following probability $\mathbb{P}\bigl[\mathbf{Y}_1 = \underline{y}_1, \dots, \mathbf{Y}_{d(n)} = \underline{y}_{d(n)}\bigr]$ as

$$\mathbb{P}\bigl[\mathbf{Y}_1 = \underline{y}_1, \dots, \mathbf{Y}_{N_0} = \underline{y}_{N_0}\bigr] \prod_{i=0}^{d(n)-1} \mathbb{P}\bigl[\mathbf{Y}_{N_i+1} = \underline{y}_{N_i+1}, \dots, \mathbf{Y}_{N_{i+1}} = \underline{y}_{N_{i+1}} \bigm| [\mathbf{W}_{N_i+1}] = ab\bigr].$$

Observe that the terms $\log \mathbb{P}\bigl[\mathbf{Y}_{N_i+1} = \cdot, \dots, \mathbf{Y}_{N_{i+1}} = \cdot \bigm| [\mathbf{W}_{N_i+1}] = ab\bigr]$, $i \in \mathbb{N}$, are i.i.d., since one can think of starting at some $\mathbf{W}_k$ with $[\mathbf{W}_k] = ab$ and stopping at the first time $l > k$ with $[\mathbf{W}_l] = ab$. By the ergodic theorem for positive recurrent Markov chains, $d(n)/n$ tends almost surely to ν_{ab}. Hence, if we consider only the subsequence where n equals one of the N_k's we obtain the following convergence for almost every trajectory $(\underline{y}_1, \underline{y}_2, \dots) \in \mathcal{W}_\pi^{\mathbb{N}}$ by classical ergodic theory:

$$\begin{aligned} & -\frac{1}{n}\log \mathbb{P}\bigl[\mathbf{Y}_1 = \underline{y}_1, \dots, \mathbf{Y}_{d(n)} = \underline{y}_{d(n)}\bigr] \\ = \; & -\frac{d(n)}{n}\frac{1}{d(n)}\Biggl[\log \mathbb{P}\bigl[\mathbf{Y}_1 = \underline{y}_1, \dots, \mathbf{Y}_{N_0} = \underline{y}_{N_0}\bigr] \\ & \quad + \sum_{i=0}^{d(n)-1} \log \mathbb{P}\bigl[\mathbf{Y}_{N_i+1} = \underline{y}_{N_i+1}, \dots, \mathbf{Y}_{N_{i+1}} = \underline{y}_{N_{i+1}} \bigm| [\mathbf{W}_{N_i+1}] = ab\bigr]\Biggr] \\ \xrightarrow{n\to\infty} \; & -\nu_{ab} \sum_{k\geq 1} \sum_{\substack{x_2,\dots,x_{k-1} \in \mathcal{W}_0: \\ [x_i] \neq ab \\ \text{for } 2 \leq i \leq k-1}} \sum_{\substack{x \in \mathcal{W}_0: \\ [x] = ab}} w(ab, x_2, \dots, x_{k-1}, x) \log \tilde{w}(ab, x_2, \dots, x_{k-1}, x). \end{aligned}$$

This proves the claim. □

8. Analyticity of Entropy

The random walk on $\mathcal{A}^*$ depends on *finitely* many parameters which are described by the transition probabilities $p(w_1,w_2)$, $w_1,w_2\in\mathcal{A}^*$ with $|w_1|\leq 2$ and $|w_2|\leq 3$; see (2.1). That is, each random walk on $\mathcal{A}^*$ can be defined via a vector $\underline{p}\in\mathbb{R}_+^{|\mathcal{B}|}$, where

$$\mathcal{B}:=\Bigl\{(w_1,w_2)\Bigm| w_1\in\mathcal{A}\cup\mathcal{A}^2\cup\{o\}, w_2\in\bigcup_{n=1}^{3}\mathcal{A}^n\cup\{o\}, \bigl||w_1|-|w_2|\bigr|\leq 1\Bigr\}.$$

In other words, the entry of $\underline{p}$ associated with the index $(w_1,w_2)\in\mathcal{B}$ describes the value of $p(w_1,w_2)$. The support $\mathrm{supp}(\underline{p})$ of $\underline{p}$ is the set of indices in $\mathcal{B}$ corresponding to non-zero entries of $\underline{p}$. Fix now any $\underline{p}_0\in\mathbb{R}_+^{|\mathcal{B}|}$ such that $\underline{p}_0$ describes a well-defined, transient random walk on $\mathcal{A}^*$, and let $\mathcal{P}(\underline{p}_0)$ be the set of vectors $\underline{p}\in\mathbb{R}^{|\mathcal{B}|}$ with support $\mathrm{supp}(\underline{p}_0)$ which allow well-defined, transient random walks on $\mathcal{A}^*$. The set $\mathcal{P}(\underline{p}_0)$ can be described by an open polygonal bounded convex set in $\mathbb{R}^d$ with some suitable $d\leq|\mathcal{B}|-1$ which depends on $\mathrm{supp}(\underline{p}_0)$; recall that $\ell>0$ if and only if $(X_n)_{n\in\mathbb{N}_0}$ is transient, and from the formula of ℓ in [7, Theorem 2.4] follows that ℓ varies continuously in $\underline{p}$, yielding that there is some open neighbourhood of $\underline{p}_0$ in $\mathbb{R}^d$ where $(X_n)_{n\in\mathbb{N}_0}$ remains still transient. We now ask whether the entropy mapping $\underline{p}\mapsto h=h_{\underline{p}}$ varies real-analytically on $\mathcal{P}(\underline{p}_0)$.

In the next subsection we will introduce a new Markov chain which is related to the last entry time process and leads under the projection $\pi(\cdot,\cdot)$ to a hidden Markov chain with same distribution as $(\mathbf{Y}_k)_{k\in\mathbb{N}}$. Afterwards we will be able to prove Theorem 2.6 in Subsection 8.2.

8.1. **Modified Last Entry Time Process.** The aim of this subsection is the construction of a Markov chain related to the last entry time process $(\mathbf{i}_k,\mathbf{W}_k)_{k\in\mathbb{N}}$ such that the transition matrix has strictly positive entries and the modified process leads under $\pi(\cdot,\cdot)$ (see (5.5)) to a hidden Markov chain with same asymptotic entropy.

Let be $ab,a_1b_1,a_2b_2\in\mathcal{A}^2$, and let $C_{j_{i,1}}$ be the first cone of type j in the covering of $C(a_1b_1)$ with $\tau(C(a_1b_1))=i$ and let $C_{j_{k,l}}$ be the l-th subcone of type j in the covering of $C(a_2b_2)$ with $\tau(C(a_2b_2))=k$. Assume that $y_0\in\partial C_{j_{k,l}}$ with $[y_0]=ab$. Since $C_{j_{i,1}}$ and $C_{j_{k,l}}$ are isomorphic, there is some unique $\bar{y}_0^{[i,j,ab]}\in\mathcal{A}^*$ such that $\bar{y}_0^{[i,j,ab]}ab\in\partial C_{j_{i,1}}$; see Section 4.1. In the following we will sometimes omit the superindex $[i,j,ab]$ and use the notation $\bar{y}_0=\bar{y}_0^{[i,j,ab]}$ for describing this replacement.

For $i,j\in\mathcal{I}$ and $ab\in\mathcal{A}^2$ with $\tau(C(ab))=j$, we write

$$\#\{j_{s,t}\mid s\neq i,ab\}:=\Bigl|\bigl\{(j_{s,t},w)\in\mathcal{W}\bigm|[w]=ab, s\in\mathcal{I}\setminus\{i\}, 1\leq t\leq n(s,j)\bigr\}\Bigr|.$$

It is not hard to see that $\#\{j_{s,t}\mid s\neq i,a_1b_1\}=\#\{j_{s,t}\mid s\neq i,a_2b_2\}$ if $\tau(C(a_1b_1))=\tau(C(a_2b_2))$ but this will not be relevant for our proofs, so we omit further explanations. Let be $(i_{k,l},x),(j_{m,n},y)\in\mathcal{W}$ with $[y]=ab\in\mathcal{A}^2$. This implies $\tau(C(x))=i$ and $y^{[i,j,ab]}\in\partial C_{j_{i,1}}$, where $C_{j_{i,1}}$ is the first cone of type j in the covering of $C([x])$. Define the following transition probabilities on $\mathcal{W}$:

$$\hat{q}\big((i_{k,l},x),(j_{m,n},y)\big) := \begin{cases} \frac{1}{\#\{j_{s,t}\mid s\neq i, ab\}+1}\frac{\xi([y])}{\xi([x])}\mathbb{L}(x,y), & \text{if } m=i\wedge n=1,\\ \frac{\xi([y])}{\xi([x])}\mathbb{L}(x,y), & \text{if } m=i\wedge n\geq 2,\\ \frac{1}{\#\{j_{s,t}\mid s\neq i, ab\}+1}\frac{\xi([y])}{\xi([x])}\mathbb{L}(x,\bar{y}^{[i,j,ab]}ab), & \text{if } m\neq i. \end{cases}$$

It is easy to see that these transition probabilities define a Markov chain (inherited from the Markov chain $(\mathbf{i}_k,\mathbf{W}_k)_{k\in\mathbb{N}}$): in the case $m=i\wedge n\geq 2$ we just have

$$\hat{q}\big((i_{k,l},x),(j_{m,n},y)\big) = \mathbb{P}\big[(\mathbf{i}_2,\mathbf{W}_2)=(j_{m,n},y)\mid(\mathbf{i}_1,\mathbf{W}_1)=(i_{k,l},x)\big];$$

otherwise we have, for $(j_{i,1},y)\in\mathcal{W}$,

$$\begin{aligned}&\hat{q}\big((i_{k,l},x),(j_{i,1},y)\big) + \sum_{\substack{(j_{s,t},w)\in\mathcal{W}:\\ s\neq i,[w]=ab,\\ 1\leq t\leq n(s,j)}}\hat{q}\big((i_{k,l},x),(j_{s,t},w)\big)\\ =\;& \mathbb{P}\big[(\mathbf{i}_2,\mathbf{W}_2)=(j_{i,1},y)\mid(\mathbf{i}_2,\mathbf{W}_2)=(i_{k,l},x)\big]\end{aligned}$$

since $y=\bar{y}^{[i,j,ab]}ab$ by definition. In other words, each step from $(i_{k,l},x)$ to $(j_{m,n},y)$ either behaves according to (5.4) (case $m=i$ and $n\geq 2$) or the step from $(i_{k,l},x)$ to $(j_{i,1},y)$ (when seen as a step of the process $(\mathbf{i}_k,\mathbf{W}_k)_{k\in\mathbb{N}}$) is split up into different equally likely steps $(i_{k,l},x)$ to $(j_{m,n},\bar{y}ab)$ with $m\neq i$ or $m=i\wedge n=1$. Observe that the transitions depend only on $[x]$ in the first argument of $\hat{q}(\cdot,\cdot)$. By Proposition 5.4, the transition matrix $\widehat{Q}=\big(\hat{q}((i_{k,l},x),(j_{m,n},y))\big)$ is stochastic and governs a positive recurrent, aperiodic Markov chain $(\mathbf{t}_k,\mathbf{x}_k)_{k\in\mathbb{N}}$. In particular, $\widehat{Q}$ has strictly positive entries. The initial distribution $\hat{\mu}_1$ of $(\mathbf{t}_1,\mathbf{x}_1)$ is defined as

$$\hat{\mu}_1(i_{m,n},x) := \mathbb{P}[(\mathbf{i}_1,\mathbf{W}_1)=(i_{m,n},x)]>0$$

for $(i_{m,n},x)\in\mathcal{W}$.

The process $\big((\mathbf{t}_k,\mathbf{x}_k),(\mathbf{t}_{k+1},\mathbf{x}_{k+1})\big)_{k\in\mathbb{N}}$ is again a positive recurrent, aperiodic Markov chain whose transition matrix is denoted by $\widehat{Q}_2$ (arising from $\widehat{Q}$). We now define a new hidden Markov chain $(\mathbf{Z}_k)_{k\in\mathbb{N}}$ by

$$\mathbf{Z}_k := \pi\big((\mathbf{t}_k,\mathbf{x}_k),(\mathbf{t}_{k+1},\mathbf{x}_{k+1})\big).$$

Observe that at this point the second branch in the definition of π in (5.5) comes into play for the definition of $\mathbf{Z}_k$. The crucial point is the following proposition:

Proposition 8.1. *For all* $(s^{(1)},t^{(1)}),\dots,(s^{(n)},t^{(n)})\in\mathcal{W}_\pi$,

$$\mathbb{P}\big[\mathbf{Y}_1=(s^{(1)},t^{(1)}),\dots,\mathbf{Y}_n=(s^{(n)},t^{(n)})\big]=\mathbb{P}\big[\mathbf{Z}_1=(s^{(1)},t^{(1)}),\dots,\mathbf{Z}_n=(s^{(n)},t^{(n)})\big].$$

Since the proof of this proposition consists of an long induction with tedious calculations we omit it at this place and give it in Appendix C.

The statement of the last proposition can be formulated in other words: the process governed by $\widehat{Q}$ can be seen as a last entry time process, where one has more subcones to enter (namely, the subcones of indices $j_{k,l},k\neq i$, when being currently in a cone of type i), but the projection π (in particular due to the second branch in its definition in (5.5)) folds the process down to the same hidden Markov chain $(\mathbf{Y}_k)_{k\in\mathbb{N}}$ in terms of probability. With Propositions 5.6 and 8.1 we immediately obtain:

Corollary 8.2. *For almost every realisation* $\big((s^{(1)},t^{(1)}),(s^{(2)},t^{(2)}),\dots\big)\in\mathcal{W}_\pi^{\mathbb{N}}$,

$$H(\mathbf{Y})=\lim_{n\to\infty}-\frac{1}{n}\log\mathbb{P}\big[\mathbf{Z}_1=(s^{(1)},t^{(1)}),\dots,\mathbf{Z}_n=(s^{(n)},t^{(n)})\big].$$

□

The important difference between the underlying Markov chains $\big((\mathbf{t}_k,\mathbf{x}_k),(\mathbf{t}_{k+1},\mathbf{x}_{k+1})\big)_{k\in\mathbb{N}}$ and $\big((\mathbf{i}_k,\mathbf{W}_k),(\mathbf{i}_{k+1},\mathbf{W}_{k+1})\big)_{k\in\mathbb{N}}$ is that the transition matrix $\widehat{Q}_2$ has strictly positive entries, while this must not necessarily hold for the transition matrix of the Markov chain $\big((\mathbf{i}_k,\mathbf{W}_k),(\mathbf{i}_{k+1},\mathbf{W}_{k+1})\big)_{k\in\mathbb{N}}$. This property will be important later.

8.2. **Proof of Theorem 2.6.** The crucial point will be the following lemma:

Lemma 8.3. *The transition probabilities* $q(w_1,w_2)$, $w_1,w_2\in\mathcal{W}_0$, *vary real-analytically w.r.t.* $\underline{p}\in\mathcal{P}(\underline{p}_0)$.

Proof. In order to show that $q(w_1,w_2)$ varies real-analytically in $\underline{p}$ it suffices to show analyticity of $H(ab,c)$, $ab\in\mathcal{A}^2$, $c\in\mathcal{A}$, and $\bar{L}(ab,cde)$, $d,e\in\mathcal{A}$, due to Proposition 5.1. The function $z\mapsto H(ab,c|z)$ has radius of convergence bigger than 1, which can be easily deduced from Lemma 6.1. Thus, for $\delta>0$ small enough, we have

$$\infty>H(ab,c|1+\delta)=\sum_{n\geq 1}\mathbb{P}_{ab}[X_n=c,\forall m<n:|X_m|\geq 2](1+\delta)^n.$$

The probability $\mathbb{P}_{ab}[X_n=c,\forall m<n:|X_m|\geq 2]$ can be rewritten as

$$\sum_{\substack{n_1,\dots,n_d\geq 1:\\ n_1+\dots+n_d=n}} c(n_1,\dots,n_d)p_1^{n_1}\cdot\ldots\cdot p_d^{n_d},\quad c(n_1,\dots,n_d)\in\mathbb{N}_0,$$

where $p_1,\dots,p_d$ correspond to the non-zero entries of the vector $\underline{p}$. Therefore,

$$H(ab,c|1+\delta)=\sum_{n\geq 1}\sum_{\substack{n_1,\dots,n_d\geq 1:\\ n_1+\dots+n_d=n}} c(n_1,\dots,n_d)(p_1(1+\delta))^{n_1}\cdot\ldots\cdot(p_d(1+\delta))^{n_d}<\infty.$$

Hence, $\underline{p}$ lies in the interior of the domain of convergence of $H(ab,c|1)$ when considered as a multivariate power series in the variables of $\mathrm{supp}(\underline{p})=\{p_1,\dots,p_d\}$. This yields real-analyticity of $H(ab,c|1)$ in $\underline{p}$. Analyticity of $\xi(ab)$ follows now directly from its definition. One can show completely analogously that the functions $\bar{L}(ab,cde|1)$ vary also real-analytically in $\underline{p}$ since $\bar{L}(ab,cde|z)$ has also radius of convergence bigger than 1, which can also be easily deduced from Lemma 6.1. This proves the statement of the lemma. □

Now we can prove:

Proof of Theorem 2.6. The claim follows now via the equation $h = \ell \cdot H(\mathbf{Y})/\lambda$. By Lemma 8.3, the invariant probability measure ν_0 of the process $(\mathbf{W}_k)_{k\in\mathbb{N}}$ varies real-analytically in some neighbourhood of $\underline{p}_0$, since ν_0 is the solution of a linear system of equations in terms of $q(\cdot,\cdot)$; hence, λ (given in (5.8)) varies analytically.

Moreover, the transition matrix $\widehat{Q}_2$ of the process $\big((\mathbf{t}_k,\mathbf{x}_k),(\mathbf{t}_{k+1},\mathbf{x}_{k+1})\big)_{k\in\mathbb{N}}$ has strictly positive entries. Therefore, we can apply the analyticity result for entropies of hidden Markov chains of Han and Marcus [12, Theorem 1.1] on $(\mathbf{Z}_k)_{k\in\mathbb{N}}$ and obtain together with Corollary 8.2 that $H(\mathbf{Y})$ is also real-analytic in some neighbourhood of $\underline{p}_0$; at this point it is crucial that $\widehat{Q}_2$ has strictly positive entries in order to be able to apply [12, Theorem 1.1], which was our motivation for the definition of the process $(\mathbf{t}_k,\mathbf{x}_k)_{k\in\mathbb{N}}$ and $(\mathbf{Z}_k)_{k\in\mathbb{N}}$.

Real-analyticity of ℓ can be shown completely analogously to the proof of Lemma 8.3 with the help of the formula for ℓ given in [7, Theorem 2.4]. This finishes the proof. □

Appendix A. Remarks on Assumptions 2.1 and 2.4

A.1. **Generalization of Suffix-Irreducibility.** In this section we make a discussion on Assumption 2.4, where we show how to relax this condition in some way and that it cannot be dropped completely. First, recall that suffix-irreducibility leads to the fact that the process $(\mathbf{W}_k)_{k\in\mathbb{N}}$ is irreducible. One can weaken the asssumption of suffix-irreducibility to the assumption that

$$\mathbb{P}[\forall n\in\mathbb{N}: |X_n|\geq |w| \mid X_0=w] > 0 \quad \forall w\in\mathcal{L}, \tag{A.1}$$

or equivalently that $H(ab,c|1) < 1$ for all $a,b,c\in\mathcal{A}$. This means that, for every $w\in\mathcal{L}$, there is some $ab\in\mathcal{A}^2$ such that

$$\mathbb{P}[\exists n\in\mathbb{N}: [X_n]=ab, \forall k\leq n: |X_k|\geq |w| \mid X_0=w] > 0 \ \text{ and } \ H(ab,\cdot|1)<1.$$

In this case the process $(\mathbf{W}_k)_{k\in\mathbb{N}}$ is not necessarily irreducible any more, but it still has a finite state space. Let $C_1,\dots,C_r$ be the essential classes of the state space of $(\mathbf{W}_k)_{k\in\mathbb{N}}$. Then $(\mathbf{W}_k)_{k\in\mathbb{N}}$ will almost surely take only values in one of these classes up to finitely many exemptions for small $k\in\mathbb{N}$; the class depends then on the concrete realization. If we condition on the fact that $(\mathbf{W}_k)_{k\in\mathbb{N}}$ will finally enter the class C_i, then – on this event – the entropy rate h_i and the drift ℓ_i can be calculated as shown in the irreducible case and as in [7]: we just have to replace $(\mathbf{W}_k)_{k\in\mathbb{N}}$ by $(\mathbf{W}_{T+k})_{k\in\mathbb{N}}$, where T is the smallest index with $\tau(\mathbf{W}_T)\in C_i$. The overall entropy rate and drift are then given by

$$\begin{aligned} h &= \lim_{n\to\infty}\frac{1}{n}\mathbb{E}[-\log\pi(X_n)] = \sum_{i=1}^r h_i\cdot\mathbb{P}[(\mathbf{W}_k)_{k\in\mathbb{N}} \text{ finally enters } C_i],\\ \ell &= \lim_{n\to\infty}\frac{1}{n}\mathbb{E}[|X_n|] = \sum_{i=1}^r \ell_i\cdot\mathbb{P}[(\mathbf{W}_k)_{k\in\mathbb{N}} \text{ finally enters } C_i]. \end{aligned}$$

Since the probabilities $\mathbb{P}[(\mathbf{W}_k)_{k\in\mathbb{N}}$ finally enters $C_i]$ are the solutions of a finite system of linear equations with coefficients $q(\cdot,\cdot)$, they vary also analytically. Hence, condition (A.1) also implies our result on analyticity of the entropy.

If the property (A.1) does not hold, then the random walk may take some long deviations between the last entry times $\mathbf{e}_{k-1}$ and $\mathbf{e}_k$ such that $\mathbb{E}[\mathbf{e}_k - \mathbf{e}_{k-1}] = \infty$; see Example A.1 below. One can show that, in the case of infinite expectation, this leads to $\lim_{n\to\infty} k/\mathbf{e}_k = 0$, implying $\liminf_{n\to\infty} |X_n|/n = 0$; an analogous statement is shown in [9], where the proof can be adapted easily to the present context. This allows no conclusion on the entropy with our techniques, since $l(X_n) = -\log L(o, X_n|1)$ can not be compared with $-\log \pi_n(X_n)$ any more as it was done in the proof of Proposition 5.10. But we underline that this setting with deviations of expected infinite length constitutes a degenerate case.

Example A.1. *Let be $\mathcal{A} = \{a, b, c, d\}$ and set*

$$\begin{aligned}
&p(o, a_1) = p(a_1, o) = \frac{1}{4}\ \forall a_1 \in \{a, b, c\},\ p(o, d) = \frac{1}{4},\ p(d, o) = \frac{1}{2},\\
&p(a_1, a_1a_2) = p(a_1a_3, a_1) = \frac{1}{4}\ \forall a_1 \in \{a, b, c\}, a_2 \in \mathcal{A}\setminus\{a_1\}, a_3 \in \mathcal{A}\setminus\{a_1, d\},\\
&p(a_1a_2, a_1a_2a_3) = \frac{1}{4}\ \forall a_1, a_2 \in \{a, b, c\}, a_1 \neq a_2, \forall a_3 \in \mathcal{A}\setminus\{a_2\},\\
&p(ad, add) = p(bd, bdd) = p(cd, cdd) = \frac{1}{2},\\
&p(d, dd) = p(dd, ddd) = p(dd, d) = \frac{1}{2},\ p(ad, a) = p(bd, b) = p(cd, c) = \frac{1}{2}.
\end{aligned}$$

The associated graph $\mathcal{G}$ can be identified as follows: the vertex set is given by $\mathbb{T}_3 \times \mathbb{N}_0$, where $\mathbb{T}_3 = (\mathbb{Z}/2\mathbb{Z}) * (\mathbb{Z}/2\mathbb{Z}) * (\mathbb{Z}/2\mathbb{Z}) = \langle a, b, c \mid a^2 = b^2 = c^2 = 1\rangle$, and the adjacency relation is defined via $(a_1 \dots a_k, m) \sim (b_1 \dots b_l, n)$ if and only if

$$\begin{cases}
a_1 \dots a_k = b_1 \dots b_l \wedge |m-n| = 1 & \textit{or}\\
m = n = 0 \wedge k = l+1 \wedge a_1 \dots a_{k-1} = b_1 \dots b_l \wedge a_k \neq a_{k-1} & \textit{or}\\
m = n = 0 \wedge k+1 = l \wedge a_1 \dots a_k = b_1 \dots b_{l-1} \wedge b_l \neq b_{l-1}.
\end{cases}$$

The graph $\mathcal{G}$ can be visualized as follows: take a homogeneous tree of degree 3, where the vertices are described by words over $\{a, b, c\}$ such that two consecutive letters are different; attach to each vertex a half-line $\mathbb{N}$, where the steps on the half-line are made with equal probability of $\frac{1}{2}$; the vertices $(w, 0)$ correspond to the vertices of the tree and one chooses with equal probability of $\frac{1}{4}$ one of the four neighbour vertices for the next step. This implies that the random walk will stay only for some finite time in each half-line before making a step in the tree part of $\mathcal{G}$. Moreover, it is not hard to see that the random walk converges to some infinite word over the subalphabet $\{a, b, c\}$. But it is well-known that the random walk needs in expectation infinite time to leave one of the halflines, that is, the expected time for reaching "a" when starting at "ad" is infinite. This implies that $\mathbb{E}[\mathbf{e}_k - \mathbf{e}_{k-1}] = \infty$.

A.2. **Weak Symmetry Assumption.** The purpose for introducing the weak symmetry assumption is that the random walk becomes irreducible and that the cones become strongly connected subgraphs. A weaker but still sufficient condition is given as follows: if $w_0 \in \mathcal{L}$ and $w_1, w_2 \in C(w_0)$ with

$$\mathbb{P}[\exists n \in \mathbb{N} : X_n = w_2, \forall m \leq n : X_m \in C(w_0) \mid X_0 = w_1] > 0$$

then $\mathbb{P}[\exists n \in \mathbb{N} : X_n = w_1, \forall m \leq n : X_m \in C(w_0) \mid X_0 = w_2] > 0$. Under this weaker condition the random walk still remains irreducible and the Green function's radius of convergence R is strictly bigger than 1. Also, the cones remain strongly connected and $C(w) = C(w')$ if $w' \in \partial C(w)$.

If this connectedness of cones is not satisfied then the definition of cones and coverings of cones by subcones gets more complicated. In that case the coverings depend on the boundary point from which one constructs the covering yielding coverings by possibly non-disjoint subcones. In particular, Lemma 4.2 does not necessarily hold. This would lead to a more detailed and complicated case distinction in order to get coverings by disjoint subcones. Since there will be no additional gain and the involving techniques remain the same we used weak symmetry for ease of presentation.

Appendix B. Switching from the K-dependent Case to the Blocked Letter Language

In this section we make a discussion on the transition from the K-dependent case (that is, the transition probabilities depend on the last K letters and between two steps of the random walk only the last K letters may be replaced by a word of length of at most $2K$) to the blocked letter language (that is, blocking words of length of at most K to new single letters such that we are in the situation defined via (2.1)). In the *K-dependent case* the general transition probabilities have the form

$$\mathbb{P}\big[X_{n+1} = wy \mid X_n = wx\big] = p(x, y), \tag{B.1}$$

where $w, x, y \in \mathcal{A}^*$ with x being a word consisting of K letters and y being a word consisting of at most $2K$ letters.

Obviously, if the K-dependent random walk is weakly symmetric then the random walk on the blocked letter language is weakly symmetric, too. Suffix-irreducibility in the K-dependent case means that, for all $w \in \mathcal{L}$ and every $w_0 \in \mathcal{A}^K$, the random walk starting at w has positive probability to visit some word ending with w_0 by only passing through words in $\mathcal{A}_{\geq |w|}$. However, suffix-irreducibility in the K-dependent case does, in general, not necessarily yield suffix-irreducibility of the blocked letter language. But as already explained in Appendix A.1 suffix irreducibility can be relaxed by the assumption (A.1), and the blocked letter language inherits this assumption from the K-dependent case.

Finally, we want to discuss the cases when the K-dependent random walk is expanding or not. Define cones in the K-dependent case as at the beginning of Subsection 4.1. For any $w \in \mathcal{A}^*$, denote by $[w]_K$ the last K letters. Two cones $C(w_1)$ and $C(w_2)$, $w_1, w_2 \in \mathcal{A}^*$ are then isomorphic if $C([w_1]_K) = C([w_2]_K)$. The same properties of cones and coverings (that is, nestedness or disjointness of cones, construction of coverings of cones by subcones, etc.) from Section 4 can be transferred to the K-dependent case analogously. If the graph $\mathcal{G}$ is *not* expanding in the K-dependent case then one can show analogously as in Subsection 4.2.2 that the random walk converges to one out of finitely many deterministic infinite words. In the following we will show that blocked letter language random walk is expanding if $\mathcal{G}$ is expanding in the K-dependent case. Recall that X_∞ is the infinite limiting random word of our K-dependent random walk.

Lemma B.1. *If the K-dependent random walk is expanding then the support of X_∞ is infinite.*

Proof. Assume that X_∞ has finite support. Choose $N \in \mathbb{N}$ large enough such that each connected component of $\mathcal{G} \setminus \{w \in \mathcal{L} \mid |w| < N\}$ (that is, remove from $\mathcal{G}$ all vertices $w \in \mathcal{L}$ with $|w| < N$ and their adjacent edges) contains in its closure only one point of the support of X_∞. Take any of these connected components and denote it by C, and take any $w_0 \in C$ with $|w_0| = N$. Then $\mathbb{P}[\forall n \geq 1 : |X_n| \geq |w_0| \mid X_0 = w_0] > 0$. Since each cone contains at least two proper subcones, we can find disjoint subcones $C(w_1), C(w_2)$ of $C(w_0)$ such that $w_1, w_2 \in \mathcal{L}$ with $|w_1| = |w_2| > |w_0| + K$. Due to condition (A.1) we have $\mathbb{P}[\forall n \geq 1 : |X_n| \geq |w_i| \mid X_0 = w_i] > 0$ for each $i \in \{1, 2\}$. We remark that this follows also from suffix-irreducibility. That is, if the random walk escapes to infinity inside $C(w_0)$ then it can escape to infinity via the cone $C(w_1)$ or via the cone $C(w_2)$, which is disjoint from $C(w_1)$. Thus, we have found two different boundary points of X_∞, which lie in the closure of C, a contradiction to our choice of N and C. Consequently, the support of X_∞ cannot be finite. □

Now we get:

Corollary B.2. *If the K-dependent random walk is expanding then the associated blocked letter language random walk is also expanding.*

Proof. Assume that the blocked letter language random walk is *not* expanding. Denote by $X_\infty^{(B)}$ the infinite limiting word w.r.t. the blocked letter language. Then $X_\infty^{(B)}$ is quasi-deterministic, that is, its support is a finite subset of $\mathcal{A}_B^{\mathbb{N}}$, where $\mathcal{A}_B$ is the blocked letter language alphabet. But this yields that X_∞ has also finite support in $\mathcal{A}^{\mathbb{N}}$, and this in turn implies by the previous lemma that the K-dependent case cannot be expanding. □

Hence, concerning the property "expanding" we have shown that there is no gain or loss when switching from K-dependent random walks to the blocked letter language random walk.

Appendix C. Proofs

In this section we give the missing proofs of some lemmas and propositions, which we omitted earlier for sake of better readability.

Proof of Lemma 4.1.
Let be $w_1 = a_1 \dots a_m$, $w_2 = b_1 \dots b_n \in \mathcal{A}^*_{\geq 2}$ with $a_1, \dots, a_m, b_1, \dots, b_n \in \mathcal{A}$ such that $C(w_1)$ and $C(w_2)$ are isomorphic.

Proof of (1): since $C(w_1)$ and $C(w_2)$ are isomorphic we have $C([w_1]) = C([w_2])$, and thus $[w_1] = a_{m-1}a_m \in C([w_1]) = C([w_2])$. Hence, there is a path $\langle [w_2], u_1, \dots, u_k, a_{m-1}a_m \rangle$ through words $u_1, \dots, u_k \in \mathcal{A}^*_{\geq 2}$. If $w' = a_1 \dots a_{m-2}\bar{w} \in C(w_1)$ with $\bar{w} \in \mathcal{A}^*_{\geq 2}$ then there is a path $\langle w_1, w_1', \dots, w_l', w' \rangle$ through words $w_1', \dots, w_l' \in \mathcal{A}^*_{\geq |w_1|}$. This yields that w_i' has the form $w_i' = a_1 \dots a_{m-2} w_i''$ with some $w_i'' \in \mathcal{A}^*_{\geq 2}$, that is, the path $\langle a_{m-1}a_m, w_1'', \dots, w_l'', \bar{w}\rangle$ has positive probability to be performed. But this implies that

$$\langle w_2 = b_1 \dots b_{n-2}[w_2], b_1 \dots b_{n-2}u_1, \dots, b_1 \dots b_{n-2}u_k, b_1 \dots b_{n-2}a_{m-1}a_m, \\ b_1 \dots b_{n-2}w_1'', \dots, b_1 \dots b_{n-2}w_l'', b_1 \dots b_{n-2}\bar{w}\rangle$$

is a path through words in $\mathcal{A}^*_{\geq |w_2|}$, that is, $b_1 \dots b_{n-2}\bar{w} \in C(w_2)$. Thus, φ is well-defined.

Since any $w \in C(w_1)$ and its image $\varphi(w)$ differ only by different (constant) prefixes the mapping φ is obviously a bijection. Moreover, if $w = a_1 \dots a_{m-2}c_1 \dots c_k \in C(w_1)$ with $c_1, \dots, c_k \in \mathcal{A}$, $k \geq 2$, and $\hat{w} = a_1 \dots a_{m-2}c_1 \dots c_{k-2}w' \in C(w_1)$ with $w' \in \mathcal{A}^*$, $1 \leq |w'| \leq 3$, and $(k-2) + |w'| \geq 2$ (otherwise $\hat{w} \notin C(w_1)$), then

$$p(w, \hat{w}) = p(c_{k-1}c_k, w') = p(b_1 \dots b_{n-2}c_1 \dots c_k, b_1 \dots b_{n-2}c_1 \dots c_{k-2}w') = p\big(\varphi(w), \varphi(\hat{w})\big).$$

This yields (1).

Proof of (2): this follows directly from (1) by the bijection φ and the fact that the adjacency relation is given through positive single-step transition probabilities. Hence, $C(w_1)$ and $C(w_2)$ are isomorphic as subgraphs of $\mathcal{G}$. □

Proof of Lemma 4.2.
Let be $w_1, w_2 \in \mathcal{A}^*_{\geq 2}$. W.l.o.g. assume that $|w_1| \leq |w_2|$. Moreover, assume that the cones $C(w_1)$ and $C(w_2)$ are not nested in each other and that $C(w_1) \cap C(w_2) \neq \emptyset$. Let be $w_0 \in C(w_1) \cap C(w_2)$. Then there is a path $\langle w_1, w_1', \dots, w_k', w_0\rangle$ through words $w_1', \dots, w_k' \in \mathcal{A}^*_{\geq |w_1|}$ and there is a path $\langle w_2, w_1'', \dots, w_l'', w_0\rangle$ through words $w_1'', \dots, w_l'' \in \mathcal{A}^*_{\geq |w_2|} \subseteq \mathcal{A}^*_{\geq |w_1|}$. By weak symmetry, there is a path $\langle w_1, w_1', \dots, w_k', w_0, w_l'', \dots, w_1'', w_2\rangle$ through words in $\mathcal{A}^*_{\geq |w_1|}$, and hence $w_2 \in C(w_1)$ which in turn implies $C(w_2) \subseteq C(w_1)$, a contradiction. This yields the first part of the lemma.

In order to prove the second part assume that $|w_1| = |w_2|$ and w.l.o.g. $C(w_1) \subseteq C(w_2)$. It remains to show that we have then $C(w_1) = C(w_2)$. Since $w_1 \in C(w_2)$ there is a path $\langle w_2, \bar{w}_1, \dots, \bar{w}_m, w_1\rangle$ through words $\bar{w}_1, \dots, \bar{w}_m \in \mathcal{A}^*_{\geq |w_2|}$. If $w \in C(w_2)$ then there is a path $\langle w_2, \hat{w}_1, \dots, \hat{w}_n, w\rangle$ through words $\hat{w}_1, \dots, \hat{w}_n \in \mathcal{A}^*_{\geq |w_2|}$. Thus, there is a path

$$\langle w_1, \bar{w}_m, \dots, \bar{w}_1, w_2, \hat{w}_1, \dots, \hat{w}_n, w\rangle$$

though words in $\mathcal{A}^*_{\geq |w_2|} = \mathcal{A}^*_{\geq |w_1|}$. Hence, $C(w_2) \subseteq C(w_1)$ which yields $C(w_2) = C(w_1)$. □

For the next proof we need the following properties: if $a_1b_1, a_2b_2 \in \mathcal{A}^2$ with $\tau(C(a_1b_1)) = \tau(C(a_2b_2))$ then we have $C(a_1b_1) = C(a_2b_2)$ (see Lemma 4.2) and therefore $a_2b_2 \in C(a_1b_1)$. In this case we also have $\mathbb{L}(a_1b_1, w) > 0$ for $w \in \mathcal{A}^*_{\geq 3}$ if and only if $\mathbb{L}(a_2b_2, w) > 0$. This follows from the simple fact that $a_2b_2 \in C(a_1b_1)$ implies that there are paths from a_1b_1 to a_2b_2 (and vice versa) through words in $\mathcal{A}^*_{\geq 2}$.

Proof of Lemma 5.2.
By definition, we obviously have $\mathrm{supp}(\mathbb{P}[\mathbf{W}_1 = \cdot]) = \mathcal{W}_0$. For $k > 1$ we show both inclusions. Let be $y \in \mathcal{W}_0$. Then there are $w_0 \in \mathcal{A}^*$ and $ab \in \mathcal{A}^2$ with $w_0ab \in \bigcup_{j=1}^{n_0} \partial C_j^{(0)}$ and

$w_0y \in \mathcal{S}(w_0ab)$ and

$$\begin{aligned}\mathbb{P}[\mathbf{W}_0 = w_0ab, \mathbf{W}_1 = y] &= \sum_{w' \in \mathcal{L}\setminus C(w_0ab)} G(o,w') \cdot p(w', w_0ab) \cdot \mathbb{L}(w_0ab, w_0y) \cdot \xi([y]) \\ &= \sum_{w' \in \mathcal{L}\setminus C(w_0ab)} G(o,w') \cdot p(w', w_0ab) \cdot \mathbb{L}(ab, y) \cdot \xi([y]) > 0.\end{aligned}$$

Take now any $\bar{w}\bar{a}\bar{b} \in \mathrm{supp}(\mathbb{P}[X_{\mathbf{e}_{k-2}} = \cdot])$. Since the covering of every cone contains subcones of all different types, the cone $C(\bar{w}\bar{a}\bar{b})$ has in its covering a cone of type $\tau(C(ab))$. Hence, there are $w_k \in \mathcal{A}^*$, $a_kb_k \in \mathcal{A}^2$ with $\bar{w}w_ka_kb_k \in \mathcal{S}(\bar{w}\bar{a}\bar{b})$, $\tau(C(a_kb_k)) = \tau(C(ab))$ and $m_k \in \mathbb{N}$ such that $p^{(m_k)}(o, \bar{w}w_ka_kb_k) > 0$. Thus,

$$\begin{aligned}\mathbb{P}[\mathbf{W}_k = y] &\geq \mathbb{P}[X_{\mathbf{e}_{k-1}} = \bar{w}w_ka_kb_k, \mathbf{W}_k = y] \\ &= \sum_{w' \in \mathcal{L}\setminus C(\bar{w}w_ka_kb_k)} G(o,w') \cdot p(w', \bar{w}w_ka_kb_k) \cdot \mathbb{L}(\bar{w}w_ka_kb_k, \bar{w}w_ky) \cdot \xi([y]) \\ &= \sum_{w' \in \mathcal{L}\setminus C(\bar{w}w_ka_kb_k)} G(o,w') \cdot p(w', \bar{w}w_ka_kb_k) \cdot \mathbb{L}(a_kb_k, y) \cdot \xi([y]).\end{aligned}$$

By the remark before the lemma, we have $\mathbb{L}(a_kb_k, y) > 0$ and therefore $\mathbb{P}[\mathbf{W}_k = y] > 0$, yielding $\mathcal{W}_0 \subseteq \mathrm{supp}(\mathbb{P}[\mathbf{W}_k = \cdot])$.

For the other direction, take any $y \in \mathrm{supp}(\mathbb{P}[\mathbf{W}_k = \cdot])$. Then there is some $w_{k-1}ab \in \mathcal{L}$ such that

$$\begin{aligned}0 &< \mathbb{P}[X_{\mathbf{e}_{k-1}} = w_{k-1}ab, X_{\mathbf{e}_k} = w_{k-1}y] \\ &= \sum_{w' \in \mathcal{L}\setminus C(w_{k-1}ab)} G(o,w') \cdot p(w', w_{k-1}ab) \cdot \mathbb{L}(w_{k-1}ab, w_{k-1}y) \cdot \xi([y]).\end{aligned}$$

In particular, $\mathbb{L}(ab, y) > 0$. Since the initial covering of $\mathcal{L}$ contains a cone of type $\tau(C(ab))$ there are $w_0 \in \mathcal{A}^*$, $a_0b_0 \in \mathcal{A}^2$ and some $m \in \mathbb{N}$ such that $w_0a_0b_0 \in \bigcup_{i=1}^{n_0} \partial C_i^{(0)}$, $\tau(C(a_0b_0)) = \tau(C(ab))$ and $p^{(m)}(o, w_0a_0b_0) > 0$. Observe again that $\mathbb{L}(a_0b_0, y) > 0$ by the remark before the lemma. Therefore,

$$\begin{aligned}\mathbb{P}[\mathbf{W}_1 = y] &\geq \mathbb{P}[\mathbf{W}_0 = w_0a_0b_0, \mathbf{W}_1 = y] = \mathbb{P}[X_{\mathbf{e}_0} = w_0a_0b_0, \mathbf{W}_1 = y] \\ &= \sum_{w' \in \mathcal{L}\setminus C(w_0a_0b_0)} G(o,w') \cdot p(w', w_0a_0b_0) \cdot \mathbb{L}(w_0a_0b_0, w_0y) \cdot \xi([y]) \\ &= \sum_{w' \in \mathcal{L}\setminus C(w_0a_0b_0)} G(o,w') \cdot p(w', w_0a_0b_0) \cdot \mathbb{L}(a_0b_0, y) \cdot \xi([y]) > 0.\end{aligned}$$

This yields $\mathrm{supp}(\mathbb{P}[\mathbf{W}_k = \cdot]) \subseteq \mathrm{supp}(\mathbb{P}[\mathbf{W}_1 = \cdot]) = \mathcal{W}_0$ and the claim of the lemma follows. □

Proof of Proposition 5.4.
It remains to show that the support of each $(\mathbf{i}_k, \mathbf{W}_k)$ equals $\mathcal{W}$ and that $\big((\mathbf{i}_k, \mathbf{W}_k)\big)_{k\in\mathbb{N}}$ is positive recurrent and aperiodic.

First, we show that $\mathrm{supp}(\mathbb{P}[(\mathbf{i}_k,\mathbf{W}_k)=\cdot])=\mathcal{W}$ for $k\geq 1$. For this purpose, let be $(j_{i,n},x)\in\mathrm{supp}(\mathbb{P}[(\mathbf{i}_k,\mathbf{W}_k)=\cdot])$. Then there is some $w_{k-1}a_{k-1}b_{k-1}\in\mathcal{L}$ with

$$\begin{aligned}&\mathbb{P}[X_{\mathbf{e}_{k-1}}=w_{k-1}a_{k-1}b_{k-1},\mathbf{W}_k=x]\\ =&\sum_{w'\in\mathcal{L}\setminus C(w_{k-1}a_{k-1}b_{k-1})} G(o,w')p(w',w_{k-1}a_{k-1}b_{k-1})\mathbb{L}(a_{k-1}b_{k-1},x)\xi([x])>0,\end{aligned}$$

$\tau(C(a_{k-1}b_{k-1}))=i$ and $C(x)$ being the n-th subcone of type j in the covering of the cone $C(a_{k-1}b_{k-1})$. If $k=1$ then $(j_{i,n},x)\in\mathcal{W}$. In the case $k>1$ take any $w_0a_0b_0\in\mathcal{L}$ with $\mathbb{P}[\mathbf{W}_0=w_0a_0b_0]>0$ and $\tau(C(w_0a_0b_0))=i$. Since $a_{k-1}b_{k-1}\in C(a_0b_0)$ we also have $\mathbb{L}(a_0b_0,x)>0$ since $\mathbb{L}(a_{k-1}b_{k-1},x)>0$ (recall the remark before Lemma 5.2). Then:

$$\mathbb{P}[\mathbf{W}_0=w_0a_0b_0,\mathbf{W}_1=x]=\sum_{w'\in\mathcal{L}\setminus C(w_0a_0b_0)} G(o,w')p(w',w_0a_0b_0)\mathbb{L}(a_0b_0,x)\xi([x])>0,$$

yielding $(j_{i,n},x)\in\mathcal{W}$.

For the other inclusion, let be $(j_{i,n},x)\in\mathcal{W}$. Then there is some $w_0a_0b_0\in\mathcal{L}$ with

$$\mathbb{P}[\mathbf{W}_0=w_0a_0b_0,\mathbf{W}_1=x]=\sum_{w'\in\mathcal{L}\setminus C(w_0a_0b_0)} G(o,w')p(w',w_0a_0b_0)\mathbb{L}(a_0b_0,x)\xi([x])>0,$$

$\tau(C(a_0b_0))=i$ and $C(x)$ being the n-th subcone of type j in the covering of $C(a_0b_0)$. If $k=1$ then $(j_{i,n},x)\in\mathrm{supp}(\mathbb{P}[(\mathbf{i}_1,\mathbf{W}_1)=\cdot])$. In the case $k>1$ take any $w_{k-2}a_{k-2}b_{k-2}\in\mathcal{L}$ with $\mathbb{P}[X_{\mathbf{e}_{k-2}}=w_{k-2}a_{k-2}b_{k-2}]>0$. Then $C(w_{k-2}a_{k-2}b_{k-2})$ has in its covering a subcone $C(w_{k-1}a_{k-1}b_{k-1})$ of type i. Since $a_{k-1}b_{k-1}\in C(a_0b_0)$ we have $\mathbb{L}(a_{k-1}b_{k-1},x)>0$ due to $\mathbb{L}(a_0b_0,x)>0$ (once again recall the remark before Lemma 5.2) and $C(x)$ is the n-th subcone of type j in the covering of $C(a_{k-1}b_{k-1})=C(a_0b_0)$. Hence,

$$\begin{aligned}&\mathbb{P}[(\mathbf{i}_k,\mathbf{W}_k)=(j_{i,n},x)]\geq\mathbb{P}[X_{\mathbf{e}_{k-1}}=w_{k-1}a_{k-1}b_{k-1},X_{\mathbf{e}_k}=w_{k-1}x]\\ \geq&\sum_{w'\in\mathcal{L}\setminus C(w_{k-1}a_{k-1}b_{k-1})} G(o,w')p(w',w_{k-1}a_{k-1}b_{k-1})\mathbb{L}(a_{k-1}b_{k-1},x)\xi([x])>0,\end{aligned}$$

yielding $\mathcal{W}\subseteq\mathrm{supp}(\mathbb{P}[(\mathbf{i}_k,\mathbf{W}_k)=\cdot])$, and therefore $\mathcal{W}=\mathrm{supp}(\mathbb{P}[(\mathbf{i}_k,\mathbf{W}_k)=\cdot])$.

The next task is to show irreducibility, which implies positive recurrence due to finiteness of $\mathcal{W}$. Let be $(i_{m,n},w_1)$, $(j_{s,t},w_2)\in\mathcal{W}$. Take any $\bar{w}\in\mathcal{W}_0$ such that $q(w_1,\bar{w})>0$ and $\tau(C(\bar{w}))=s$, which exists by construction of coverings. Then $w_2\in\partial C_{j_{s,t}}([\bar{w}])$, that is, $C(w_2)$ is the t-th subcone of type j in the covering of $C([\bar{w}])$, yielding $q(\bar{w},w_2)>0$. Hence,

$$\begin{aligned}&\mathbb{P}[(\mathbf{i}_3,\mathbf{W}_3)=(j_{s,t},w_2)\mid(\mathbf{i}_1,\mathbf{W}_1)=(i_{m,n},w_1)]\\ \geq\ &\mathbb{P}[\mathbf{W}_3=w_2,\mathbf{W}_2=\bar{w}\mid(\mathbf{i}_1,\mathbf{W}_1)=(i_{m,n},w_1)]\\ =\ &q(w_1,\bar{w})\cdot q(\bar{w},w_2)>0.\end{aligned}\tag{C.1}$$

Here, we used the fact that $\mathbf{i}_3=j_{s,t}$ is uniquely determined by $w_1,\bar{w},w_2$ and that this probability does not depend on m and n. This yields irreducbility of the process $\big((\mathbf{i}_k,\mathbf{W}_k)\big)_{k\in\mathbb{N}}$.

It follows that the period of the process is at most 2. In order to see aperiodicity, take any $\underline{w_1},\underline{w_2}\in\mathcal{W}$ with $\mathbb{P}[(\mathbf{i}_2,\mathbf{W}_2)=\underline{w_2}\mid(\mathbf{i}_1,\mathbf{W}_1)=\underline{w_1}]>0$. Then we get analogously to (C.1):

$$\begin{aligned}&\mathbb{P}[(\mathbf{i}_4,\mathbf{W}_4)=\underline{w_1},(\mathbf{i}_2,\mathbf{W}_2)=\underline{w_2}\mid(\mathbf{i}_1,\mathbf{W}_1)=\underline{w_1}]\\ =\ &\mathbb{P}[(\mathbf{i}_2,\mathbf{W}_2)=\underline{w_2}\mid(\mathbf{i}_1,\mathbf{W}_1)=\underline{w_1}]\cdot\mathbb{P}[(\mathbf{i}_4,\mathbf{W}_4)=\underline{w_1}\mid(\mathbf{i}_2,\mathbf{W}_2)=\underline{w_2}]>0.\end{aligned}$$

That is, the period of the process is 1. This finishes the proof. □

The following lemma was used in the proof of Proposition 5.8:

Lemma C.1. *For almost every realisation* $(x_1, \underline{y}_1, \underline{y}_2, \dots)$ *of* $(X_{\mathbf{e}_1}, \mathbf{Y}_1, \mathbf{Y}_2, \dots)$,

$$H(\mathbf{Y}) = \lim_{k\to\infty} -\frac{1}{k} \log \mathbb{P}\big[C(X_{\mathbf{e}_1}) = C(x_1), \mathbf{Y}_1 = \underline{y}_1, \dots, \mathbf{Y}_k = \underline{y}_k\big].$$

Proof. We recall the notation from the proof of Proposition 5.8: let be $k \in \mathbb{N}$ and assume for the moment that $\mathbf{W}_l = y_l a_l b_l$, where $y_l \in \mathcal{A}^* \setminus \{o\}$ and $a_l b_l \in \mathcal{A}^2$ for $0 \leq l \leq k$. That is, $X_{\mathbf{e}_l} = y_0 y_1 \dots y_l a_l b_l$. We write $\mathbf{Y}_1 = (j, t^{(1)})$, where $j = \tau(C(a_1 b_1))$, and $\mathbf{Y}_l = (s^{(l)}, t^{(l)})$ for $2 \leq l \leq k$, where the values of $s^{(2)}, \dots, s^{(k-1)}$ and $t^{(1)}, \dots, t^{(k-1)}$ are determined by the values of $\mathbf{W}_l = y_l a_l b_l$. Vice versa, given $X_{\mathbf{e}_1}$ the values of $s^{(2)}, \dots, s^{(k-1)}$ and $t^{(1)}, \dots, t^{(k-1)}$ determine uniquely the cones $C(y_l a_l b_l)$: indeed, $X_{\mathbf{e}_1}$ and $t^{(1)}$ determine uniquely $C(X_{\mathbf{e}_2})$ and therefore also $C(\mathbf{W}_2) = C(y_2 a_2 b_2)$; inductively, given $C(X_{\mathbf{e}_l})$ of type $s^{(l)}$ then $t^{(l)}$ determines uniquely $C(X_{\mathbf{e}_{l+1}})$ and $C(\mathbf{W}_{l+1}) = C(y_{l+1} a_{l+1} b_{l+1})$. We mark it by $(*)$ when we make use of this "transition".

Recall that the covering of $\mathcal{L}$ consists of n_0 subcones $C_i^{(0)}$, $1 \leq i \leq n_0$. Each $C_i^{(0)}$ has again a covering consisting of $n(\tau(C_i^{(0)}), j)$ subcones of type j. We enumerate all these subcones of type j by $C_{j,k}^{(1)}$ with $1 \leq k \leq N_j := \sum_{i=1}^{n_0} n(\tau(C_i^{(0)}), j)$, that is, we enumerate all subcones of type j which appear in the coverings of all $C_i^{(0)}$, $1 \leq i \leq n_0$.

Since $\mathcal{W}_0$ is finite, there is some constant $c > 0$ such that

$$c \cdot \mathbb{P}[X_{\mathbf{e}_1} = x] \leq \mathbb{P}[X_{\mathbf{e}_1} = y]$$

for all $x, y \in \bigcup_{k=1}^{N_j} \partial C_{j,k}^{(1)} \subseteq \mathrm{supp}(\mathbb{P}[X_{\mathbf{e}_1} = \cdot])$.

In the following we will show that $\mathbb{P}\big[C(X_{\mathbf{e}_1}) = C(x_1), \mathbf{Y}_1 = \underline{y}_1, \dots, \mathbf{Y}_k = \underline{y}_k\big]$ is comparable with $\mathbb{P}\big[\mathbf{Y}_1 = \underline{y}_1, \dots, \mathbf{Y}_k = \underline{y}_k\big]$, which proves the claim. First, we have for $k \geq 2$:

$$\begin{aligned}
& N_j \cdot \mathbb{P}\big[X_{\mathbf{e}_1} \in C(y_0 y_1 a_1 b_1), \mathbf{Y}_1 = (j, t^{(1)}), \mathbf{Y}_2 = (s^{(2)}, t^{(2)}), \dots, \mathbf{Y}_{k-1} = (s^{(k-1)}, t^{(k-1)})\big] \\
\stackrel{(*)}{=}\ & N_j \cdot \sum_{x \in \partial C(y_0 y_1 a_1 b_1)} \sum_{\substack{w_2, \dots, w_k \in \mathcal{W}_0: \\ w_i \in \partial C(y_i a_i b_i) \\ \text{for all } 2 \leq i \leq k}} \mathbb{P}[X_{\mathbf{e}_1} = x, X_{\mathbf{e}_2} = y_0 y_1 w_2, \dots, X_{\mathbf{e}_k} = y_0 \dots y_{k-1} w_k] \\
=\ & N_j \cdot \sum_{\substack{x \in \partial C(y_0 y_1 a_1 b_1); \\ w_2, \dots, w_k \in \mathcal{W}_0: \\ w_i \in \partial C(y_i a_i b_i) \\ \text{for all } 2 \leq i \leq k}} \mathbb{P}[X_{\mathbf{e}_1} = x] \mathbb{P}[X_{\mathbf{e}_2} = y_0 y_1 w_2, \dots, X_{\mathbf{e}_k} = y_0 \dots y_{k-1} w_k \mid X_{\mathbf{e}_1} = x]
\end{aligned}$$

$$= N_j \cdot \sum_{x\in\partial C(y_0y_1a_1b_1)} \sum_{\substack{w_2,\dots,w_k\in\mathcal{W}_0:\\ w_i\in\partial C(y_ia_ib_i)\\ \text{for all } 2\leq i\leq k}} \mathbb{P}[X_{\mathbf{e}_1}=x]q(y_1[x],w_2)\prod_{i=3}^k q(w_{i-1},w_i)$$

$$= \sum_{l=1}^{N_j} \sum_{x\in\partial C(y_0y_1a_1b_1)} \sum_{\substack{w_2,\dots,w_k\in\mathcal{W}_0:\\ w_i\in\partial C(y_ia_ib_i)\\ \text{for all } 2\leq i\leq k}} \mathbb{P}[X_{\mathbf{e}_1}=x]\mathbb{P}\big[\mathbf{W}_2=w_2\big|[X_{\mathbf{e}_1}]=[x]\big]\prod_{i=3}^k q(w_{i-1},w_i).$$

For a moment, let be $\partial C(y_0y_1a_1b_1)=\{y_0y_1c_1d_1,\dots,y_0y_1c_\kappa d_\kappa\}$. Then for all $l\in\{1,\dots,N_j\}$ there is some $v_l\in\mathcal{A}^*$ such that $\partial C_{j,l}^{(1)}=\{v_lc_1d_1,\dots,v_lc_\kappa d_\kappa\}$. Therefore, for every $x\in\partial C(y_0y_1a_1b_1)$ and each $l\in\{1,\dots,N_j\}$ there is exactly one $\hat{x}_l\in\partial C_{j,l}^{(1)}$ with $[\hat{x}_l]=[x]$, $\mathbb{P}[X_{\mathbf{e}_1}=x]\geq c\cdot\mathbb{P}[X_{\mathbf{e}_1}=\hat{x}_l]$ and $\mathbb{P}\big[\mathbf{W}_2=w_2\big|[X_{\mathbf{e}_1}]=[x]\big]=\mathbb{P}\big[\mathbf{W}_2=w_2\big|[X_{\mathbf{e}_1}]=[\hat{x}_l]\big]$ for all $w_2\in\mathcal{W}_0$. The last equation follows from the fact that the probabilities depend on $X_{\mathbf{e}_1}$ only by its last two letters $[X_{\mathbf{e}_1}]$ in the condition. We write $\hat{x}_l$ for this mapping $(x,l)\mapsto\hat{x}_l$. Hence,

$$N_j\cdot\mathbb{P}\begin{bmatrix} X_{\mathbf{e}_1}\in C(y_0y_1a_1b_1),\mathbf{Y}_1=(j,t^{(1)}),\\ \mathbf{Y}_2=(s^{(2)},t^{(2)}),\dots,\mathbf{Y}_{k-1}=(s^{(k-1)},t^{(k-1)})\end{bmatrix}$$

$$\geq \sum_{l=1}^{N_j} \sum_{x\in\partial C(y_0y_1a_1b_1)} \sum_{\substack{w_2,\dots,w_k\in\mathcal{W}_0:\\ w_i\in\partial C(y_ia_ib_i)\\ \text{for all } 2\leq i\leq k}} c\mathbb{P}[X_{\mathbf{e}_1}=\hat{x}_l]\mathbb{P}\big[\mathbf{W}_2=w_2\mid[X_{\mathbf{e}_1}]=[\hat{x}_1]\big]\prod_{i=3}^k q(w_{i-1},w_i)$$

$$= \sum_{l=1}^{N_j} \sum_{w\in\partial C_{j,l}^{(1)}} \sum_{\substack{w_2,\dots,w_k\in\mathcal{W}_0:\\ w_i\in\partial C(y_ia_ib_i)\\ \text{for all } 2\leq i\leq k}} c\cdot\mathbb{P}[X_{\mathbf{e}_1}=w]\cdot\mathbb{P}\big[\mathbf{W}_2=w_2\big|[X_{\mathbf{e}_1}]=[w]\big]\cdot\prod_{i=3}^k q(w_{i-1},w_i)$$

$$= c\cdot\mathbb{P}\big[\mathbf{Y}_1=(j,t^{(1)}),\mathbf{Y}_2=(s^{(2)},t^{(2)}),\dots,\mathbf{Y}_{k-1}=(s^{(k-1)},t^{(k-1)})\big].$$

Vice versa, we obviously have

$$\begin{aligned} &\mathbb{P}\big[X_{\mathbf{e}_1}\in C(y_0y_1a_1b_1),\mathbf{Y}_1=(j,t^{(1)}),\mathbf{Y}_2=(s^{(2)},t^{(2)}),\dots,\mathbf{Y}_{k-1}=(s^{(k-1)},t^{(k-1)})\big]\\ \leq\ &\mathbb{P}\big[\mathbf{Y}_1=(j,t^{(1)}),\mathbf{Y}_2=(s^{(2)},t^{(2)}),\dots,\mathbf{Y}_{k-1}=(s^{(k-1)},t^{(k-1)})\big].\end{aligned}$$

This proves the claim. □

Proof of Proposition 8.1. Let be $(s^{(1)},t^{(1)}),\dots,(s^{(n)},t^{(n)})\in\mathcal{W}_\pi$. We prove the claim by induction on n. First, let be $j,s\in\mathcal{I}$ and $t^{(1)}=j_m$ with $2\leq m\leq n(s,j)$, and let $a_0b_0,ab\in\mathcal{A}^2$ with $\tau(C(a_0b_0))=s$ and $\tau(C(ab))=j$. If $C_{j,m}$ is the m-th cone of type j in the covering of $C(a_0b_0)$ then there is a unique word $\bar{x}_0=\bar{x}_0^{[s,j,m,ab]}\in\mathcal{A}^*$ with $\bar{x}_0ab\in\partial C_{j,m}$.

With this notation we get:

$$\begin{aligned}\mathbb{P}\bigl[\mathbf{Y}_1=(s,j_m),[\mathbf{W}_2]=ab\bigr] &= \sum_{(u_{k,l},x)\in\mathcal{W}:u=s}\mathbb{P}[(\mathbf{i}_1,\mathbf{W}_1)=(s_{k,l},x)]\cdot q(x,\bar{x}_0ab)\\ &= \sum_{(u_{k,l},x)\in\mathcal{W}:u=s}\hat{\mu}_1(s_{k,l},x)\hat{q}\bigl((s_{k,l},x),(j_{s,m},\bar{x}_0ab)\bigr)\\ &= \mathbb{P}\bigl[\mathbf{Z}_1=(s,j_m),[\mathbf{x}_2]=ab\bigr].\end{aligned}$$

Now we turn to the case $t^{(1)}=j_1$. Once again, if $C_{j,1}$ is the first cone of type j in the covering of $C(a_0b_0)$ then there is some unique $\bar{x}_0=\bar{x}_0^{[s,j,1,ab]}\in\mathcal{A}^*$ with $\bar{x}_0ab\in\partial C_{j,1}$. We get:

$$\begin{aligned}&\mathbb{P}\bigl[\mathbf{Z}_1=(s,j_1),[\mathbf{x}_2]=ab\bigr]\\ =&\sum_{(u_{k,l},x)\in\mathcal{W}:u=s}\hat{\mu}_1(s_{k,l},x)\Bigl[\hat{q}\bigl((s_{k,l},x),(j_{s,1},\bar{x}_0ab)\bigr)+\sum_{\substack{(t_{p,q},y)\in\mathcal{W}:\\ t=j,p\neq s,[y]=ab}}\hat{q}\bigl((s_{k,l},x),(j_{p,q},y)\bigr)\Bigr]\\ =&\sum_{\substack{(u_{k,l},x)\in\mathcal{W}:\\ u=s}}\hat{\mu}_1(s_{k,l},x)\Biggl[\frac{q\bigl((s_{k,l},x),(j_{s,1},\bar{x}_0ab)\bigr)}{\#\{t_{\kappa_1,\kappa_2}\mid\kappa_1\neq s,ab\}+1}+\sum_{\substack{(t_{p,q},y)\in\mathcal{W}:\\ t=j,p\neq s,\\ [y]=ab}}\frac{q\bigl((s_{k,l},x),(j_{s,1},\bar{x}_0ab)\bigr)}{\#\{t_{\kappa_1,\kappa_2}\mid\kappa_1\neq s,ab\}+1}\Biggr]\\ =&\sum_{(u_{k,l},x)\in\mathcal{W}:u=s}\mathbb{P}\bigl[(\mathbf{i}_1,\mathbf{W}_1)=(s_{k,l},x)\bigr]\cdot q(x,\bar{x}_0ab)=\mathbb{P}\bigl[\mathbf{Y}_1=(s,j_1),[\mathbf{W}_2]=ab\bigr].\end{aligned}$$

Now, in both cases we obtain

$$\begin{aligned}\mathbb{P}\bigl[\mathbf{Z}_1=(s,t^{(1)})\bigr] &= \sum_{ab\in\mathcal{A}^2}\mathbb{P}\bigl[\mathbf{Z}_1=(s,t^{(1)}),[\mathbf{x}_2]=ab\bigr]\\ &= \sum_{ab\in\mathcal{A}^2}\mathbb{P}\bigl[\mathbf{Y}_1=(s,t^{(1)}),[\mathbf{W}_2]=ab\bigr]=\mathbb{P}\bigl[\mathbf{Y}_1=(s,t^{(1)})\bigr].\end{aligned}$$

We now perform the induction step where we will use the induction assumption

$$\begin{aligned}&\mathbb{P}\bigl[\mathbf{Y}_1=(s^{(1)},t^{(1)}),\dots,\mathbf{Y}_n=(s^{(n)},t^{(n)}),[\mathbf{W}_{n+1}]=ab\bigr]\\ =&\ \mathbb{P}\bigl[\mathbf{Z}_1=(s^{(1)},t^{(1)}),\dots,\mathbf{Z}_n=(s^{(n)},t^{(n)}),[\mathbf{x}_{n+1}]=ab\bigr].\end{aligned}\tag{C.2}$$

First, consider the case $(s^{(n+1)},t^{(n+1)})=(s,j_m)$ with $s,j\in\mathcal{I}$ and $2\leq m\leq n(s,j)$. This implies that $\mathbf{t}_{n+1}$ has the form $s_{*,*}$ and $\mathbf{t}_{n+2}=j_{s,m}$. Let $C_{j,m}$ be the m-th cone of type j in the covering of $C(a_0b_0)$, where $a_0b_0\in\mathcal{A}^2$ with $\tau(C(a_0b_0))=s$. If $ab\in\mathcal{A}^2$ with $\tau(C(ab))=j$ then there is some unique $\bar{x}_0=\bar{x}_0^{[s,j,m,ab]}\in\mathcal{A}^*$ with $\bar{x}_0ab\in\partial C_{j,m}$. In this case we obtain:

$$\begin{aligned}
&\mathbb{P}\big[\mathbf{Z}_1=(s^{(1)},t^{(1)}),\dots,\mathbf{Z}_{n+1}=(s^{(n+1)},j_m),[\mathbf{x}_{n+1}]=a_0b_0,[\mathbf{x}_{n+2}]=ab\big]\\
=&\sum_{\substack{(u_{k,l},w_0)\in\mathcal{W}:\\ u=s,[w_0]=a_0b_0}}\mathbb{P}\big[\mathbf{Z}_1=(s^{(1)},t^{(1)}),\dots,\mathbf{Z}_n=(s^{(n)},t^{(n)}),\mathbf{t}_{k+1}=u_{k,l},\mathbf{x}_{n+1}=w_0\big]\\
&\quad\cdot\hat{q}\big((s_{k,l},w_0),(j_{s,m},\bar{x}_0ab)\big)\\
=&\ \mathbb{P}\big[\mathbf{Z}_1=(s^{(1)},t^{(1)}),\dots,\mathbf{Z}_n=(s^{(n)},t^{(n)}),[\mathbf{x}_{n+1}]=a_0b_0\big]\frac{\xi(ab)}{\xi(a_0b_0)}\mathbb{L}(a_0b_0,\bar{x}_0ab)\\
=&\ \mathbb{P}\big[\mathbf{Y}_1=(s^{(1)},t^{(1)}),\dots,\mathbf{Y}_n=(s^{(n)},t^{(n)}),[\mathbf{W}_{n+1}]=a_0b_0\big]\frac{\xi(ab)}{\xi(a_0b_0)}\mathbb{L}(a_0b_0,\bar{x}_0ab)\\
=&\ \mathbb{P}\big[\mathbf{Y}_1=(s^{(1)},t^{(1)}),\dots,\mathbf{Y}_{n+1}=(s^{(n+1)},j_m),[\mathbf{W}_{n+1}]=a_0b_0,[\mathbf{W}_{n+2}]=ab\big].
\end{aligned}$$

Now we turn to the case $(s^{(n+1)},t^{(n+1)})=(s,j_1)$. This implies again that $\mathbf{t}_{n+1}$ has the form $s_{*,*}$. Once again, if $C_{j,1}$ is the first cone of type j in the covering of $C(a_0b_0)$ (of type s) then there is some unique $\bar{x}_0=\bar{x}_0^{[s,j,1,ab]}\in\mathcal{A}^*$ with $\bar{x}_0ab\in\partial C_{j,1}$. We get by distinguishing whether $t^{(n+1)}=j_1$ arises from $\mathbf{t}_{n+2}=j_{s,1}$ or $\mathbf{t}_{n+2}=j_{k,l}$ with $k\neq s$:

$$\begin{aligned}
&\mathbb{P}\big[\mathbf{Z}_1=(s^{(1)},t^{(1)}),\dots,\mathbf{Z}_{n+1}=(s^{(n+1)},j_1),[\mathbf{x}_{n+1}]=a_0b_0,[\mathbf{x}_{n+2}]=ab\big]\\
=&\sum_{\substack{(u_{p,q},w_0)\in\mathcal{W}:\\ u=s,[w_0]=a_0b_0}}\mathbb{P}\big[\mathbf{Z}_1=(s^{(1)},t^{(1)}),\dots,\mathbf{Z}_n=(s^{(n)},t^{(n)}),\mathbf{t}_{n+1}=u_{p,q},\mathbf{x}_{n+1}=w_0\big]\\
&\quad\cdot\Big(\hat{q}\big((s_{p,q},w_0),(j_{s,1},\bar{x}_0ab)\big)+\sum_{\substack{(t_{k,l},y)\in\mathcal{W}:\\ t=j,k\neq s,[y]=ab}}\hat{q}\big((s_{p,q},w_0),(j_{k,l},y)\big)\Big)\\
=&\ \mathbb{P}\big[\mathbf{Z}_1=(s^{(1)},t^{(1)}),\dots,\mathbf{Z}_n=(s^{(n)},t^{(n)}),[\mathbf{x}_{n+1}]=a_0b_0\big]\\
&\quad\cdot\left[\frac{\xi(ab)}{\xi(a_0b_0)}\frac{\mathbb{L}(a_0b_0,\bar{x}_0ab)}{\#\{j_{k,l}\mid k\neq s,ab\}+1}+\sum_{\substack{(t_{k,l},y)\in\mathcal{W}:\\ t=j,k\neq s,\\ [y]=ab}}\frac{\xi(ab)}{\xi(a_0b_0)}\frac{\mathbb{L}(a_0b_0,\bar{x}_0ab)}{\#\{j_{\kappa_1,\kappa_2}\mid\kappa_1\neq s,ab\}+1}\right]\\
=&\ \mathbb{P}\big[\mathbf{Z}_1=(s^{(1)},t^{(1)}),\dots,\mathbf{Z}_n=(s^{(n)},t^{(n)}),[\mathbf{x}_{n+1}]=a_0b_0\big]\frac{\xi(ab)}{\xi(a_0b_0)}\mathbb{L}(a_0b_0,\bar{x}_0ab)\\
=&\ \mathbb{P}\big[\mathbf{Y}_1=(s^{(1)},t^{(1)}),\dots,\mathbf{Y}_n=(s^{(n)},t^{(n)}),[\mathbf{W}_{n+1}]=a_0b_0\big]\frac{\xi(ab)}{\xi(a_0b_0)}\mathbb{L}(a_0b_0,\bar{x}_0ab)\\
=&\ \mathbb{P}\big[\mathbf{Y}_1=(s^{(1)},t^{(1)}),\dots,\mathbf{Y}_{n+1}=(s^{(n+1)},j_1),[\mathbf{W}_{n+1}]=a_0b_0,[\mathbf{W}_{n+2}]=ab\big].
\end{aligned}$$

Hence,

$$\begin{aligned}
& \mathbb{P}\big[\mathbf{Z}_1=(s^{(1)},t^{(1)}),\dots,\mathbf{Z}_{n+1}=(s^{(n+1)},t^{(n+1)}),[\mathbf{x}_{n+2}]=ab\big]\\
= & \sum_{a_0b_0\in\mathcal{A}^2}\mathbb{P}\big[\mathbf{Z}_1=(s^{(1)},t^{(1)}),\dots,\mathbf{Z}_{n+1}=(s^{(n+1)},t^{(n+1)}),[\mathbf{x}_{n+1}]=a_0b_0,[\mathbf{x}_{n+2}]=ab\big]\\
= & \sum_{a_0b_0\in\mathcal{A}^2}\mathbb{P}\big[\mathbf{Y}_1=(s^{(1)},t^{(1)}),\dots,\mathbf{Y}_{n+1}=(s^{(n+1)},t^{(n+1)}),[\mathbf{W}_{n+1}]=a_0b_0,[\mathbf{W}_{n+2}]=ab\big]\\
= & \mathbb{P}\big[\mathbf{Y}_1=(s^{(1)},t^{(1)}),\dots,\mathbf{Y}_{n+1}=(s^{(n+1)},t^{(n+1)}),[\mathbf{W}_{n+2}]=ab\big].
\end{aligned}$$

This proves Equation (C.2) for all $n\in\mathbb{N}$, all $ab\in\mathcal{A}^2$ and all $(s^{(1)},t^{(1)}),\dots,(s^{(n)},t^{(n)})\in\mathcal{W}_\pi$. Finally, we obtain:

$$\begin{aligned}
& \mathbb{P}\big[\mathbf{Z}_1=(s^{(1)},t^{(1)}),\dots,\mathbf{Z}_{n+1}=(s^{(n+1)},t^{(n+1)})\big]\\
= & \sum_{ab\in\mathcal{A}^2}\mathbb{P}\big[\mathbf{Z}_1=(s^{(1)},t^{(1)}),\dots,\mathbf{Z}_{n+1}=(s^{(n+1)},t^{(n+1)}),[\mathbf{x}_{n+2}]=ab\big]\\
= & \sum_{ab\in\mathcal{A}^2}\mathbb{P}\big[\mathbf{Y}_1=(s^{(1)},t^{(1)}),\dots,\mathbf{Y}_{n+1}=(s^{(n+1)},t^{(n+1)}),[\mathbf{W}_{n+2}]=ab\big]\\
= & \mathbb{P}\big[\mathbf{Y}_1=(s^{(1)},t^{(1)}),\dots,\mathbf{Y}_{n+1}=(s^{(n+1)},t^{(n+1)})\big].
\end{aligned}$$

This finishes the proof. □

References

[1] A. Avez. Entropie des groupes de type fini. *C.R.Acad.Sci.Paris Sér. A-B*, 275:A1363–A1366, 1972.

[2] I. Benjamini and Y. Peres. Tree-indexed random walks on groups and first passage percolation. *Probab. Theory Related Fields*, 98(1):91–112, 1994.

[3] S. Blachère, P. Haïssinsky, and P. Mathieu. Asymptotic entropy and Green speed for random walks on countable groups. *Ann. Probab.*, 36(3):1134–1152, 2008.

[4] T. Ceccherini-Silberstein and W. Woess. Context-free pairs of groups I: Context-free pairs and graphs. *European J. Combin.*, 33:1449–1466, 2012.

[5] T. Cover and J. Thomas. *Elements of Information Theory.* Wiley & Sons, 2nd edition, 2006.

[6] L. Gilch and F. Ledrappier. Regularity of the drift and entropy of random walks on groups. *Publ. Mat. Urug.*, 14:147–158, 2013.

[7] L. A. Gilch. Rate of escape of random walks on regular languages and free products by amalgamation of finite groups. *Discrete Math. Theor. Comput. Sci. Proc.*, pages 405–420, 2008.

[8] L. A. Gilch. Asymptotic entropy of random walks on free products. *Electron. J. Probab.*, 16:76–105, 2011.

[9] L. A. Gilch and S. Müller. Random walks on directed covers of graphs. *J. Theoret. Probab.*, 24(1):118–149, 2009.

[10] S. Gouëzel. Martin boundary of measures with infinite support in hyperbolic groups. *Ann. Probab*, to appear.

[11] P. Haïssinsky, P. Mathieu, and S. Müller. Renewal theory for random walks on cocompact Fuchsian groups. *Preprint*, 2012.

[12] G. Han and B. Marcus. Analyticity of entropy rate of hidden Markov chains. *IEEE Trans. Inform. Theory*, 52(12):5251–5266, 2006.

[13] J. Hopcraft and J. Ullman. *Introduction to Automata Theory, Languages, and Computation.* Addison-Wesley, Reading MA, 1979.

[14] V. Kaimanovich and W. Woess. Boundary and entropy of space homogeneous Markov chains. *Ann. Probab.*, 30:323–363, 2002.

[15] J. Kingman. The ergodic theory of subadditive processes. *J. Roy. Statist. Soc., Ser. B*, 30:499–510, 1968.

[16] S. Lalley. Random walks on regular languages and algebraic systems of generating functions. *Algebraic Methods in Statistics and Probability, Contemp. Math.*, 287(201–230), 2000.

[17] F. Ledrappier. Analyticity of the entropy for some random walks. *Groups Geom. Dyn.*, 6:317–333, 2012.

[18] F. Ledrappier. Regularity of the entropy for random walks on hyperbolic groups. *Ann. Probab.*, 41(5):3582–3605, 2013.

[19] V. Malyshev. Stabilization laws in the evolution of a random string. *Problems Inform. Transmission*, 30:260–274, 1995.

[20] P. Mathieu. Differentiating the entropy of random walks on hyperbolic groups. *Ann. Probab.*, 43(1):166–187, 2015.

[21] D. Muller and P. Schupp. Groups, the theory of ends and context-free languages. *J. Comput. System Sci.*, 26:295–310, 1983.

[22] D. Muller and P. Schupp. The theory of ends, pushdown automata, and second-order logic. *Theoret. Comput. Sci.*, 37:51–75, 1985.

[23] W. Woess. Context-free pairs of groups. II – Cuts, tree sets, and random walks. *Discr. Math.*, 312:157–173, 2012.

Lorenz Gilch: Graz University of Technology, Steyrergasse 30/III/C304, 8010 Graz, Austria

E-mail address: Lorenz.Gilch@freenet.de

URL: http://www.math.tugraz.at/∼gilch/

Multiple Authorship

Since Publications B and C have been written in collaboration with other authors, I have to explain my contribution to these research articles.

Generally speaking, I don't think that it is possible to measure the contribution of each author in a fair way. In my community research articles are written after many fruitful discussions with the co-authors as well as with other people. These discussions are as important as finding the right idea when thinking on one's own about some mathematical problem. Therefore, it does not make sense to assign some percentage to each author since each author contributed essentially to the successfully published research articles in different ways. Furthermore, I want to mention that the authors are listed in alphabetical order as it is standard in almost all research papers in my community.

Acknowledgements

The author is grateful to Wolfgang Woess for numerous discussions on several problems and his help during the preparation of this thesis. Furthermore, I would like to thank my co-authors for many fruitful discussions and for the close amicable collaboration.