Michael KALAB, BSc

Linearly Edge-Reinforced Random Walk

MASTER'S THESIS

to achieve the university degree of
Diplom-Ingenieur

Master's degree programme:
Operations Research and Statistics

submitted to

Graz University of Technology

Supervisor
Univ.-Prof. Dipl.-Ing. Dr. rer. nat. Wolfgang WOESS

Institute of Discrete Mathematics

Graz, February 2018

AFFIDAVIT

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master's thesis.

Date

Signature

Abstract

We review known results on linearly edge-reinforced random walk ($LRRW$), a non-Markovian stochastic process on infinite graphs $G = (V, E)$. The process is shown to be a mixture of Markov chains, both in the recurrent and the general case. We use de Finetti's results on exchangeable sequences to prove existence and uniqueness of the mixing measure in the recurrent case, following the proof lines of Diaconis and Freedman. For the general case we follow Merkl and Rolles' proof on the existence of a mixing measure. Showing recurrence or transience on certain graphs is non-trivial, the existence of a phase transition is conjectured but not proved. However, in case G is a $(K + 1)$-regular tree it is well-known that a phase transition occurs at a certain parameter $a_0(K)$. On general graphs this matter is much more difficult. Angel, Crawford and Kozma proved recurrence for sufficiently small intial weights on graphs with bounded degrees and transience for sufficiently large initial weights on non-amenable graphs. We only concentrate on the first result. In the last section we state and prove some results on the mixing measure, which have been obtained in the course of the elaboration of this thesis. We consider mixing measures for different initial vertices and show that these are mutually absolutely continuous. The respecitive derivatives have a closed and simple form. We prove the uniqueness of the mixing measure on a sub-σ-algebra of the space of Markov chains. This gives reason to believe in the uniqueness of the mixing measure even in the general case while the problem of showing this remains open.

Keywords: reinforced random walk; edge-reinforced random walk; random walks on graphs; mixture of Markov chains; mixing measure; stochastic processes with reinforcement; Pólya urn; partial exchangeability; de Finetti

For a lost friend.

Contents

Chapter 1

Introduction

Imagine travelling to an unknown city and exploring it on foot. One will usually walk along the safer streets, occasionally also choose dark and dangerous alleys by coincidence. However, using a street over and over again will become routine after some time and will thus make it more likely to walk along the street again. So, when modelling the walk, we need to take the last idea into account. Indeed, the probability of using a street will increase with the number of walks on it. In this setting the term 'reinforced' may be understood as the development of some routine. The street map will be modelled as an undirected graph $G = (V, E)$. Crossings correspond to vertices $v \in V(T)$ and streets to edges $e \in E(G)$. The safety of the streets at the beginning of the walk is modelled by a vector of positive initial weights $(a_e)_{e \in E(G)}$, usually we take $a_e \equiv a$ for some $a \in \mathbb{R}^+$. The development of a routine corresponds directly to a set of non-decreasing functions $w_e(\cdot)$ on edges $e \in E(G)$. The main purpose of Section 2 is to introduce a general model for the above issue and to specify what is meant by linearly edge-reinforced random walk on G with initial weights a, in the following abbreviated by $LRRW(G, a)$. Instead of exploring a city we might also consider a salesman moving from one market place to a neighboring one every day. In this case the reinforcement does not concern the streets (neighborhood relation) but the market places (vertices). The model associated with this type is known as vertex-reinforced random walk. Since this thesis does not investigate this kind of process the term 'reinforced' will be used instead of 'edge-reinforced'.

Generally speaking, throughout the last decade there has been a lot of research concerning stochastic processes with reinforcement. The study is not only of theoretical interest. Certain bacteria modify the chemical environment in their surroundings and are therefore either more or less likely to move to the same place again. Myxobacteria produce a slime on which they are able to glide. Once edge-reinforced random walk gives hope to understand the latter. We will just introduce

the model and one result by Codling, [3], in Section 2.
Another purpose is to understand tumour-induced angiogenesis. In tumour-like tissues the growth of vessels is enhanced and thus the migration of vessel-like cells can be modelled as a reinforced stochastic process, see Levine, [12].
The development of opinions as well as evolutionary processes may be modelled taking into account some reinforcement. A very famous and well-studied model for the first is Pólya's urn. Take an urn containing a finite number of balls of different types. A ball is drawn from the urn according to some probability distribution and the content is modified according to the type of the ball. In the basic model the probability of the choice of a type is proportional to the number of balls of the type. Many generalizations of Pólya's urn have been studied assuming different distributions as well as positive real numbers of balls or uncountably many balls, see for instance Pemantle, [16, p.22–25]. The simple generalization for positive real initial ball numbers will be introduced to study linearly reinforced random walk on trees in Section 3.1. This model has a few useful properties. One of them is that the probability of a finite sequence of balls stays the same when permuting the sequence, called exchangeability. A rigorous definition of exchangeability and a generalization for the case of walks on graphs, partial exchangeability, are introduced in Section 3.1. A consequence of the first property is almost sure convergence of the urn's relative content to a Dirichlet distributed random variable. We are going to exploit this in various ways. The proof of this fact is given later in Section 3.3.
All processes mentioned above are non-Markovian, their future depends not only on the current state but on the entire history. Curiously, some of them may still be represented, not by a single Markov chain, but as an integral w.r.t. some chains.This representation, in the following called mixture, is introduced in Section 3.2. The concept of mixtures is, of course, easier assuming that a process considers only random variables defined on the same space. In this case we establish the representation via de Finetti's representation theorem on exchangeable sequences of random variables. To appeal on intuition the Hewitt-Savage 0-1-law for the exchangeable σ-algebra of sequences of i.i.d. random variables is stated and proved in advance in Section 3.3. We then give a rigorous proof of de Finetti's theorem. The representation for recurrent $LRRW$ is then constructed from the latter in Section 3.4. The theorem and its proof are due to Diaconis and Freedman, [6]. The key property is partial exchangeability. For 2-connected graphs G the class of partially exchangeable random walks coincides with the class of the linearly reinforced one. However, if G fails to be 2-connected this does not hold true. A counterexample in the case of a tree and a rigorous proof for the 2-connected case have been obtained in the course of the elaboration this thesis and may be found at the end of Section 3.1.

Without the assumption of returning infinitely often to the starting point the proof technique of Diaconis and Freedman breaks down. Merkl and Rolles, [15], used a tightness argument to prove the existence of a representation of $LRRW$ as a mixture of Markov chains even if it is transient. We will state and prove this in Section 3.6. For this purpose the usual notation of reversible Markov chains by their associated transition matrices will be cumbersome. We describe these Markov chains and introduce the new notation in chapter 3.5. In the focus of the remaining part of the thesis are results on recurrence and transience. In the case of trees a dichotomy between recurrence and transience for $LRRW$ is a minor result at the beginning of Section 4.1. For equal initial weights $a_e \equiv a$ and $(K+1)$-regular trees we give a proove the existence of a phase transition. The phase transition occurs at a certain parameter $a_0(K)$, $LRRW$ being almost surely positive recurrent if $a < a_0(K)$ and almost surely transient if $a > a_0(K)$. In this context the term 'almost sure' has to be understood w.r.t. the mixing measure μ. The material of this section is due to Pemantle, [17]. He proved the last result not only for fixed trees but Galton-Watson trees. On general graphs $LRRW$ turns out to be much more difficult. In 2012 Angel, Crawford and Kozma, [1], studied $LRRW$ on graphs with bounded vertex degree K. They proved the existence of a parameter $a_0(K)$ so that $LRRW(G, a)$ is almost surely recurrent if $a \leqslant a_0$. The assumption $a_e \equiv a$ is not needed. We will give a proof of this result in Section 4.2. Many problems concerning $LRRW$ remain unsolved. For instance, it appears only natural to believe that for all graphs $LRRW$ is either recurrent or transient for fixed initial weights. It seems that neither a rigorous proof of this nor a counterexample has been found yet. It is as well conjectured that for a fixed graph increasing a may only make the process 'more transient' in the sense that the process does not switch between recurrence and transience while increasing a.

Chapter 2

The Model

2.1 Preliminaries

On a graph $G = (V, E)$ we may define a stochastic process as follows. Let $P = p(\cdot,\cdot)$ be a row stochastic matrix on $V(G) \times V(G)$ satisfying

$$p(v, w) > 0 \Leftrightarrow \{v, w\} \in E(G).$$

We generate sequence of vertices $(X_t)_{t\in\mathbb{N}}$ by the following scheme. Start at a fixed vertex x_0. At each time step $t + 1$ we choose a follower X_{t+1} of X_t among the vertices adjacent to X_t. We choose X_{t+1} according to the transition probabilities in P_{X_t}, the X_t-th row of P. The following properties hold for the process resulting from the above.

$$\mathbb{P}(X_{t+1} = v|X_1, \ldots, X_t) = \mathbb{P}(X_{t+1} = v|X_t) \text{ and} \tag{2.1}$$
$$\mathbb{P}(X_{t+1} = v|X_t) = \mathbb{P}(X_{t+k+1} = v|X_{t+k}). \tag{2.2}$$

Processes fulfilling these properties are called discrete, time homogenous Markov chains. The term time homogenous refers to the second equation, the transition matrix does not change over time. In the following we will omit the terms discrete and time homogenous. The Markov chains we will need to represent *LRRW* fulfill two more properties. This will be shown later on in Section 3.4.

Definition 2.1.1 (Cycles, Reversibility and Irreducibility)**.** *Let G be a graph. A cycle $c = (x_0, x_1, x_2, \ldots, x_l = x_0)$ in G is a path starting and ending in the same vertex x_0 and containing each vertex at most once. Its reversal is $c^- = (x_0 = x_l, x_{l-1}, x_{l-2}, \ldots, x_0)$. We say that a Markov chain is reversible if the following holds for any cycle c, for any index $t \in \mathbb{N}$*

$$\mathbb{P}\left((X_t, X_{t+1}, \ldots, X_{t+l}) = c\right) = \mathbb{P}\left((X_t, X_{t+1}, \ldots, X_{t+l}) = c^-\right). \tag{2.3}$$

It is called irreducible if for each ordered pair (v, w) of vertices there is some $N \in \mathbb{N}$ so that the probability of visiting w from v within N steps is positive.

The above definition of reversibility is less restrictive than in most definitions that can be found in literature but will be convenient for the rest of this assignment. For non-reinforced irreducible random walk visiting a state once almost surely implies that all states are visited infinitely often almost surely. This does not hold a priori for reinforced random walk. We will therefore choose the following definitions of (positive) recurrence and transience.

Definition 2.1.2. *A state is called recurrent (transient) if it is visited infinitely (finitely) often almost surely. We call it positive recurrent if the expected return time to the initial vertex is finite. A stochastic process on a discrete state space is called (positive) recurrent (transient) if all its states are (positive) recurrent (transient).*

The following model of reinforced random walk is due to Kozma, [11].

Definition 2.1.3 (Reinforced Random Walk)**.** *Let $G = (V, E)$ an undirected locally finite connected graph and $(w_e)_{e\in E(G)} : \mathbb{N} \to (0, \infty)$. The w_e are sometimes called conductances or weights. To avoid confusion, the notion 'initial weight' denotes a positive vector a, while 'weights' is reserved for random variables. We call the $w_e(\cdot)$ reinforcement functions and a value $w_e(t)$ routine. For a vertex v let $\mathcal{N}_G(v)$ denote the set of vertices adjacent to v in G. Let X be an infinite path in G, i.e. a sequence of adjacent vertices. Let $t \in \mathbb{N}, v := X_t, e \in E(G)$ and define for any edge e*

$$N(e, t) := |\{n \in \mathbb{N} : 1 \leqslant n \leqslant t, \{X_{n-1}, X_n\} = e\}|$$

the number of transversals of edge e up to time n. Note that we do not care about the direction of transversal. Define for all $u \in \mathcal{N}_G(v)$

$$\mathbb{P}(X_{t+1} = u|X_0, X_1, \ldots, X_t) := \frac{w_{\{u,v\}}(N(\{u, v\}, t))}{\sum_{x\in\mathcal{N}_G(v)} w_{\{x,v\}}(N(\{x, v\}, t)}. \tag{2.4}$$

If X satisfies (2.4)*, we say it is reinforced random walk with reinforcement $(w_e)_{e\in E(G)}$.*

With this definition the probability of choosing an edge is proportional to the value of $w_e(N(e, t))$. We will only deal with non-decreasing functions $w_e(\cdot)$ in this assignment. Local finiteness of G is not only a simplifying condition but also ensures that the walk X is well-defined. If, for instance, there are infinitely many edges incident to a vertex v, the denominator in (2.4) need not be a finite number.

2.2 Once Reinforced Random Walk

Consider the reinforcement function

$$w_e(n) = \begin{cases} a & n = 0 \\ a+1 & \text{else.} \end{cases}$$

This simplified model was introduced by Davis, [5], due to the long absence of theoretical results for $LRRW$. While intuitively much simpler the problem turned out to be anything but easy. In April 2016 Kious and Sidoravicius, [10, p.2], were able to prove the existence of a phase transition in a on $\mathbb{Z}^d$-like trees, the first result of this kind for once reinforced random walk.

2.3 Superlinear Reinforcement

We call a reinforced random walk superlinear if the reinforcement functions w_e are non-decreasing and for all $e \in E(G)$

$$\sum_{i=0}^{\infty} \frac{1}{w_e(i)} < \infty. \tag{2.5}$$

Superlinearly reinforced random walk on a locally finite graph turns out to get stuck almost surely on one edge under some minor assumptions, the result is due to Limic and Tarres, [13]. We will not give the entire proof but only concentrate on a proof idea.

Theorem 2.3.1. *Let G be a locally finite graph, let e be the first edge traversed, let $w := w_e$ be a non-decreasing function satisfying*

$$\sum_{i=0}^{\infty} \frac{1}{w(i)} < \infty. \tag{2.6}$$

Then

$$\mathbb{P}(\textit{The process gets stuck on e for eternity.}) > 0.$$

Proof. Let $e = \{u, v\}$ be the first edge traversed and let $K := \max(\deg_G(u), \deg_G(v))$. Then

$$\mathbb{P}(\text{The process gets stuck on one edge for eternity}) \geqslant$$
$$\prod_{i=0}^{\infty} \frac{w(i)}{(K-1)w(0) + w(i)} = \prod_{i=0}^{\infty} \left(1 - \frac{(K-1)f(0)}{(K-1)f(0) + w(i)}\right) > 0$$

since all factors are strictly positive and the infinite product is a positive number if and only if the sum in (2.6) converges. □

2.4 Linear Reinforcement

We will only consider one type of function, namely $w_e(n) = a_e + n$ for all edges e for positive constants a_e. If the increment was different, say for instance $d > 0$, we would easily transform $a' = \frac{a}{d}$. By looking at (2.4) it is obvious that this transformation has no impact on the resulting process. If the w_e are increased by 1 after each transversal the sum in (2.6) does not converge. But it diverges very slowly, the partial sums are of order $\mathcal{O}(\ln(n))$. We denote linearly reinforced random walk on a graph G by $LRRW(G, a)$ where $a = (a_e)_{e \in E(G)}$ are the initial weights. We will at first not make any assumptions on the initial vertex, but just assume that **some** initial vertex x_0 is given.

2.4.1 Expected Return Time is not Finite

The following example implies that the expected return time is not finite, except for trivial cases where $|V(G)| \leqslant 3$. It is inspired by Angel, Crawford and Kozma, [1].

Example 2.4.1. *Let G be a locally finite graph, let $a_e \equiv a$. Let x_0 be the initial vertex and $e = \{x_1, x_2\}$ some edge at distance 1 from x_0. Let $K := \max\{\deg_G(x_0), \deg_G(x_1), \deg_G(x_2)\}$. For $M \in \mathbb{N}$ denote by E_M the event that $LRRW(G, a)$ moves from x_0 to x_1 and then traverses edge e M times back and forth. Of course, $\mathbb{P}(T_0 \geqslant 2M) \geqslant \mathbb{P}(E_M)$. The probability of E_M may be estimated from below by*

$$\begin{aligned} \mathbb{P}(E_M) &\geqslant \frac{1}{K} \prod_{i=0}^{M-1} \frac{2i+a}{2i+1+Ka} \prod_{i=0}^{M-1} \frac{2i+1+a}{2i+1+Ka} = \\ &= \frac{1}{K} \prod_{i=1}^{M-1} \left(1 - \frac{(K-1)a}{2i+1} + O\left(\frac{1}{i^2}\right)\right)^2 \geqslant C(K) M^{-(K-1)a} \end{aligned} \tag{2.7}$$

for some $C(K) > 0$. Now observe that for $a \leqslant \frac{1-\epsilon}{K-1}$ for some $\epsilon > 0$ we have

$$\mathbb{E}[T_0] \geqslant 2M\mathbb{P}(T_0 \geqslant 2M) \geqslant 2MC(K)M^{-1+\epsilon} = 2C(K)M^{\epsilon}$$

for some $C(K) > 0$. Thus the expected return time cannot be finite for sufficiently small a.

The situation is not much different if we drop the conditions $a \leqslant \frac{1-\epsilon}{K}$ and $a_e \equiv a$. (2.7) is increasing in a_e. Denote by $\delta_G(v)$ the set of edges incident to v. Let a' be different intial weights, let $a^* = \max\{a'_e, e \in \delta_G(x_1) \cup \delta_G(x_2)\}$. For $N \in \mathbb{N}$ we define E_N analogously to E_M. Assuming that E_N ocurs we may as well adjust the *initial weights and suppose that the process starts at x_1, still having the focus on the first visit of x_0. Choosing N sufficiently large we observe that*

$$\mathbb{P}(E_M|E_N, a') \geqslant \mathbb{P}(E_M|a)$$

holds for all $M \in \mathbb{N}$. Since the probability of E_N is positive the expected return time is infinite also in the general case.

Chapter 3

Exchangeability and Mixtures

3.1 LRRW on Trees and Pólya's Urns

Pólya's urn model works the following way. We start with an urn containing 2 balls of different colour. At each time step we draw a ball from the urn with probability proportional to the number of its kind. It is returned, followed by an extra ball of its kind. This stochastic process has many convenient properties. We are going to compare it with $LRRW$ on a star graph first.
The material of the following paragraph is due to Pemantle, [17]. Consider the star graph S_5 consisting of the center x_0 and 5 additional vertices. $LRRW(S_5, 1)$ for initial vertex x_0 may be modelled as a Pólya's urn process. It starts with an urn containing 5 balls of 5 different colours. Each colour may be understood as an edge. Again, at each time a ball of a is drawn from the urn with probability proportional to the number of its kind, i.e. an edge is chosen with probability proportional to its routine. It is returned, but now followed by **two** extra balls of its kind. This is clear viewing that an edge must be traversed back and forth by lack of other possibilities. We may, of course, choose a different initial vertex v. In this case we still take the center x_0 as the initial vertex, but initially with two balls of the colour corresponding to $\{x_0, v\}$ instead of one. For the model corresponding to $LRRW(S_5, 1)$ let $\bar{U}_n$ be the relative content at time n, i.e. a stochastic vector whose entries are the probabilities for the next colour chosen. Let B_n be the colour of the ball drawn in the n-th step.

1. $\bar{U}_n$ satisfy equation (2.1)! This follows directly from interpreting U_n as the vector of transition probabilities. Notably, (2.2) does not hold.

2. The probability of any finite sequence $(B_1, B_2, \ldots, B_n)$ is invariant w.r.t. finite permutations, i.e. for all $\rho \in \mathfrak{S}_n$.

$$\mathbb{P}(B_1 = b_1, \ldots, B_n = b_n) = \mathbb{P}(B_1 = b_{\rho(1)}, \ldots, B_n = b_{\rho(n)}). \tag{3.1}$$

Now let $T = (V, E)$ be a locally finite **tree**. We attach an urn U^v at each vertex $v \in V(T)$ containing $\deg_T(v)$ different balls (i.e. one for each edge) and perform $LRRW(T, 1)$ for a fixed initial vertex x_0. We assume that it is recurrent. Let T_n^v denote the time of the n-th visit of a vertex v. Let U_n^v denote the relative content of the urn attached at vertex v at time T_n^v. Leaving a vertex v via edge e implies returning to v via e. Observe that the edge by which v is visited first is determined by the tree structure of T and x_0. Thus we may split up $LRRW(T, 1)$ into a set of independent processes $LRRW(S_v, \cdot)$ corresponding directly to urns U^v, so for each collection of vertices $v_1, v_2, \ldots, v_k$ the sequences

$$(U_n^{v_1})_{n\in\mathbb{N}}, \ldots, (U_n^{v_k})_{n\in\mathbb{N}}$$

are jointly independent.
This model can easily be extended to $LRRW\left(T, (a_e)_{e\in E(T)}\right)$ for any initial weight vector $a > 0$ as long as recurrence holds true. The foregoing discussion and the following theorem are due to Pemantle, [17].

Theorem 3.1.1. *Suppose an urn contains u_i balls of type i at the beginning for $1 \leqslant i \leqslant m$ and that at each step* ***one*** *ball is added to the urn. Again we denote by $U_n = (U_{n,1}, \ldots, U_{n,m})$ the vector containing the amounts of balls of each colour in the n-th step. Define by*

$$\bar{U}_n := \frac{U_n}{\sum_{i=1}^m U_{n,i}}, \quad \bar{u}_i = \frac{u_i}{\sum_{i=1}^m u_i}$$

the relative number of balls after the n-th chosen ball. At each step a ball is added to the urn, with probability for being of type i equal to $\bar{U}_{n,i}$. The u_i need not be integers, this is well-defined for any positive u_i. Then $\bar{U}_n$ converges almost surely to a random variable U whose distribution has density

$$f_U(p_1, \ldots, p_m) = \frac{\Gamma(\sum_{i=1}^m u_i)}{\prod_{i=1}^m \Gamma(u_i)} \prod_{i=1}^{m-1} p_i^{u_i - 1}(1 - p_1 - \cdots - p_{m-1})^{u_m}.$$

The proof of this Theorem is delayed until the end of Section 3.2.
A process B satisfying (3.1) is called exchangeable. The definition of **partial** exchangeability is not consistent in literature. Diaconis and Freedman as well as Rolles defined partial exchangeablity by the number of directed, respectively undirected edge transversals. We will choose the latter by Rolles, [19], since we only observe the process on undirected graphs. The definition of exchangeability may be extended to random variables on general spaces, see for instance Hewitt and Savage, [8], and Diaconis, [21]. Hewitt and Savage use the term 'symmetric' instead of 'exchangeable'.

Definition 3.1.2 ((Partial) Exchangeablity). *1. Let Z be some countable state space, X be a $Z^{\mathbb{N}}$-valued random variable. We call X and its probability measure $\mathbb{P}$ exchangeable if $\mathbb{P}$ is invariant w.r.t. finite permutations, i.e. for all $n \in \mathbb{N}$ for all $\rho \in \mathfrak{S}_{n+1}$*

$$\mathbb{P}(X_0 = x_0, \ldots, X_n = x_n) = \mathbb{P}(X_0 = x_{\rho(0)}, \ldots, X_n = x_{\rho(n)}).$$

2. Let $G = (V, E)$ be a graph, X a $V(G)^{\mathbb{N}}$-valued random variable on the set of paths in G. We call two finite directed paths y, y' equivalent,

$$y \equiv y',$$

if they start in the same vertex and for each edge e, y and y' traverse e equally often. In this context e is considered to be undirected. Naturally, y and y' need to have the same length. We call X and its probability measure $\mathbb{P}$ ***partially*** *exchangeable if for all y, y', $y \equiv y'$*

$$\mathbb{P}(y) = \mathbb{P}(y').$$

In other words $\mathbb{P}$ is constant on equivalence classes w.r.t. $\equiv$.

Lemma 3.1.3. *LRRW is partially exchangeable.*

Proof. The probability of a specific path is a product of expressions of the form in (2.4). All denominators (and of course all numerators) are determined by the number of edge transversals and these are the same for equivalent paths. □

However, an exchangeable reinforced random walk need not be *LRRW*. If $|V(G)| = 2$, any reinforced random walk is partially exchangeable. But also for general trees there are partially exchangeable processes that are not *LRRW* as the following example shows. Let T be a locally finite tree, x_0 be the initial vertex. For all edges e define the reinforcement function recursively by

$$w_e(0) = a_0, \quad w_e(n) = \begin{cases} w_e(n-1) + d_1 & \text{if } n \text{ even} \\ w_e(n-1) + d_2 & \text{if } n \text{ odd} . \end{cases}$$

It is easy to see that the resulting process is partially exchangeable although it is not *LRRW*. We would need to extend the definition of *LRRW* to directed graphs. But still, each time leaving v, the total routine $\sum_{e \in \delta_G(v)} w_e(N(e,t))$ of edges incident to v must be determined by the number of times v has been visited. The following characterization of *LRRW* on 2-connected graphs has been obtained in the course of the elaboration of this thesis.

Theorem 3.1.4. *Let X be some reinforced random walk with reinforcement function $(w_e)_{e\in E(G)}$ on a graph $G = (V, E)$, $|V(G)| \geqslant 3$. If X is partially exchangeable then for all $e \in E(G)$, for all $n \in \mathbb{N}$*

$$w_e(n+2) - w_e(n) = 2d \tag{3.2}$$

for a non-negative constant d. If, additionally, e is contained in a cycle then

$$w_e(n+1) - w_e(n) = d$$

.

Proof. Let $y = (x_0, x_1, \ldots, x_k)$ be a finite path in G and let $X_0 = x_0$ almost surely. By abuse of notation let y also be the event of performing y in the first steps of X. Its probability is

$$\mathbb{P}(y) = \prod_{i=0}^{k-1} \frac{w_{\{x_i,x_{i+1}\}}(N(\{x_i, x_{i+1}\}, i))}{\sum_{x\in\mathcal{N}_G(x_i)} w_{\{x_i,x\}}(N(\{x_i, x\}, i))}. \tag{3.3}$$

We only look at denominators

$$\sum_{x\in\mathcal{N}_G(x_i)} w_{\{x_i,x\}}(N(\{x_i, x\}, i)) \tag{3.4}$$

for $x_i \in V(G)$, the numerators being determined anyway by the transition numbers. We are going to prove the first part by choosing two paths $y_1 \equiv y_2$. G contains the line $L_G = (\{x_0, x_1, x_2\}, \{\{x_0, x_1\}, \{x_1, x_2\}\})$ as an induced subgraph, eventually renaming the vertices in G. Suppose the process has already been walking along a path y of length $n - 1 \geqslant 0$ and is at x_1 right at the moment. Let $N_{ij} = N(\{x_i, x_j\}, n)$ be the number of transversals of $\{x_i, x_j\}$ up to time n. Denote by $w_{-L} = \sum_{x_j\in\mathcal{N}_G(x_0)\setminus\{x_0,x_2\}} w_{1j}(N_{1j})$ the sum of routines over all edges incident to x_1 and not in $E(L_G)$. Denote by $[y, y', \ldots]$ the concatenation of paths $y, y', \ldots$. Define

$$y_1 = (x_1, x_0, x_1, x_2, x_1) \text{ and } y_2 = (x_1, x_2, x_1, x_0, x_1).$$

Clearly $y_1 \equiv y_2$. By partial exchangeability $\mathbb{P}([y, y_1]) = \mathbb{P}([y, y_2])$. As we do not want to write the probabilitiy of $[y, y_1]$ and of $[y, y_2]$ in terms of the form in 3.3 we observe:

1. All terms that correspond to transitions within y and to the transition from the last vertex of y and x_1 are equal and cancel out.

2. All numerators on the left side appear exactly once on the right side, so all numerators cancel out.

3. The denominators corresponding to the following transitions within y_1 and y_2, respectively, are pairwise equal.
   - $x_1 \to x_0$ within y_1 and $x_1 \to x_2$ within y_2
   - $x_0 \to x_1$ within y_1 and $x_0 \to x_1$ within y_2
   - $x_2 \to x_1$ within y_1 and $x_2 \to x_1$ within y_2

Cancelling out these terms in $\mathbb{P}([y, y_1]) = \mathbb{P}([y, y_2])$ yields

$$w_{-L} + w_{12}(N_{12}) + w_{01}(N_{01} + 2) = w_{-L} + w_{12}(N_{12} + 2) + w_{01}(N_{01}) \Leftrightarrow$$
$$w_{01}(N_{01} + 2) - w_{01}(N_{01}) = w_{12}(N_{12} + 2) - w_{12}(N_{12}).$$

Since this holds true for all pairs of edges incident to each other the first claim of the theorem holds.
To prove the second part let $c = (x_0, x_1 \dots, x_k, x_{k+1} = x_0)$ be a cycle in G. W.l.o.g. the process starts in x_0 (or else let x_0 be the node in c visited first). Let y_s be a path that runs $n \geqslant 0$ times through the whole cycle and traverses the edge $\{x_{k-1}, x_k\}$ s times additionally back and forth. We denote this path by y_s. Take $y_1 = (x_0, x_1, \dots, x_k, x_0, x_k)$, $y_2 = (x_0, x_k, x_0, x_1, \dots, x_k)$. Partial exchangeability implies $\mathbb{P}([y_s, y_1]) = \mathbb{P}([y_s, y_2])$, so

$$[w_{0,k}(n) + w_{k-1,k}(2s + n + 1)] \cdot [w_{0,k}(n + 1) + w_{0,1}(n + 1)] =$$
$$[w_{0,k}(n + 1) + w_{k-1,k}(2s + n)] \cdot [w_{0,k}(n + 2) + w_{0,1}(n)]$$

Simplifying $w_{k-1,k}(2s + \cdot) =: w_s, w_{0,k} =: g, w_{0,1} = h$ yields

$$[g(n) + w_s(n + 1)] \cdot [g(n + 1) + h(n + 1)] = [g(n + 1) + w_s(n)] \cdot [g(n + 2) + h(n)]. \tag{3.5}$$

By taking $y_1 = (x_0, x_1, x_0, x_k, \dots, x_1), y_2 = (x_0, x_k, \dots, x_1, x_0, x_1)$, i.e. swapping w and h, we obtain a second equation.

$$[g(n) + h(n + 1)] \cdot [g(n + 1) + w_s(n + 1)] = [g(n) + h(n + 1)] \cdot [g(n + 2) + w_s(n)] \tag{3.6}$$

Subtracting (3.6) from (3.5), reordering and simple calculation steps lead to

$$\frac{w_s(n + 1) - h(n + 1)}{w_s(n) - h(n)} = \frac{g(n + 2) - g(n + 1)}{g(n + 1) - g(n)}$$

Now manipulating w by s finishes the proof. The left side tends to 1 for $s \to \infty$, the right side can take but two values (for n odd resp. even) and thus needs to be equal to 1. Hence $g(n + 2) - g(n + 1) = g(n + 1) - g(n)$ and the second claim holds true. $\square$

We finish this chapter with a corollary to 3.1.4.

Corollary 3.1.5. *Let $G = (V, E)$ be a locally finite connected graph. Let X be some partially exchangeable reinforced random walk on G. If G is 2-connected then X is LRRW.*

Proof. Each edge in a 2-connected graph is contained in a cycle. □

3.2 Mixtures of Markov Chains

For many results in probability theory concerning sequences of random variables it is assumed that these random variables are i.i.d. or at least pairwise independent. Already for the simplest reinforced process in this assignment, Pólya's urn model, this assumption clearly does not hold. Due to the various results following from independence a useful technique is to represent a sequence of dependent random variables by integrating over the space of sequences of i.i.d. random variables.

Definition 3.2.1 (Mixtures of i.i.d. sequences)**.** *Let $(\mathcal{X}, \tau)$ be a Polish space, let $\mathcal{B}(\mathcal{X})$ be the Borel-σ-algebra generated by sets in τ. We write $\mathcal{B}(\mathcal{X})^{\mathbb{N}} := \bigotimes_{n\in\mathbb{N}} \mathcal{B}(\mathcal{X})$ for its product-σ-algebra. Let Y be a random variable with image contained in $\mathcal{X}$ and $(Y_n)_{n\in\mathbb{N}^*}$ be a sequence of not necessarily independent $\mathcal{X}$-valued random variables, let $\mathbb{P}$ be the probability measure associated with (Y_n).*
Let $\mathcal{P} \subset \mathfrak{m}(\mathcal{X})$ be a family of probability distributions on $\mathcal{B}(\mathcal{X})$ and $\mathcal{B}(\mathcal{P})$ be the σ-algebra generated by the sets in its weak topology. (Y_n) is called mixture of i.i.d. sequences if there exists a measure μ on $\mathcal{B}(\mathcal{P})$ so that for all $A \in \mathcal{B}(\mathcal{X})^{\mathbb{N}}$

$$\mathbb{P}(A) = \int_{\mathcal{P}^1} \pi^{\mathbb{N}}(A) d\mu(\pi)$$

*where $\pi^{\mathbb{N}}$ is the product $\otimes\pi$ of countably many **identical** replicas $\pi \in \mathcal{P}$.*

In Bayesian statistics, the law of μ is usually called the prior of $\mathbb{P}$. It may be interpreted as a 'density on a density'. For a stochastic path process on a graph the analogue of π is the transition matrix of a Markov chain. For a given state v the conditional random variables

$$X_{t+1} | X_t = v$$

may be regarded as a (maybe finite) sequence of i.i.d. random variables. Markov chains are a big field of study by themselves. Therefore we introduce the notion of mixture of Markov chains. The following definition is due to Diaconis, [6].

Definition 3.2.2. *Let $G = (V, E)$ be a graph, X a $V(G)^{\mathbb{N}}$-valued random variable, $X_0 = x_0$ for some $x_0 \in V(G)$. Let $\mathbb{P}$ be the probability measure associated with X. Denote by $\mathcal{P}$ the space of stochastic matrices $P = (p(v, w))_{v,w \in V(G)}$ on $V(G) \times V(G)$ for which*

$$p(v, w) > 0 \Rightarrow \{v, w\} \in E(G). \tag{3.7}$$

We denote by $\mathcal{B}(\mathcal{P})$ the σ-algebra generated by all sets of the form

$$\{P \in \mathcal{P} | p(v, w) \in A, A \subset [0, 1], A \text{ open}, \{v, w\} \in E(G)\}. \tag{3.8}$$

X is called a mixture of Markov chains if there exists a probability measure μ on $\mathcal{B}(\mathcal{P})$ that satisfies for all $n \in \mathbb{N}$

$$\mathbb{P}(X_i = x_i, 0 \leqslant i \leqslant n) = \int_{\mathcal{P}} \prod_{i=0}^{n-1} p(x_i, x_{i+1}) d\mu(P).$$

We are typically interested in ergodic properties. In the case of Markov chains these are, due to 0-1-laws, well-understood. For mixtures these break down. Possibly, some ergodic events happen with probability 1 one some Markov chains but with probability 0 on others. Showing recurrence in case of a mixture is thus a much more difficult task.

3.3 De Finetti and the Exchangeable σ-Algebra

To make the connection between mixtures of i.i.d. random variables and Markov chains explicit some notation is required. Let $\mathcal{X}$ be a the space of finite paths starting and ending in x_0 and not visiting x_0 in between. If X is some recurrent walk we may cut it into pieces, each piece belonging to $\mathcal{X}$. Thus we refer to X as a sequence of $\mathcal{X}$-valued random variables. Due to its countability we equip $\mathcal{X}$ with the discrete topology. As a consequence the Borel-σ-algebra $\mathcal{B}(\mathcal{X})$ is simply the power set of $\mathcal{X}$. If X is partially exchangeable, the pieces are exchangeable. By Lemma 3.1.3 this holds especially true if X is *LRRW*. This property motivates the definition of the exchangeable σ-algebra.

Definition 3.3.1. *Let $\mathcal{X}, \mathcal{B}(\mathcal{X})$, $\mathbb{P}$ and $(Y_n)_{n \in \mathbb{N}}$ as in Definition 3.2.1. For $N \in \mathbb{N}^*$ define by $\mathcal{E}_N$ the σ-algebra containing all sets $E \in \mathcal{B}(\mathcal{X})^{\mathbb{N}}$ for which*

$$\mathbb{P}((Y_n) \in E) = \mathbb{P}((Y_{\rho(n)}) \in E)$$

for all permutations ρ on $\mathbb{N}^*$ that leave all but $1, 2, \ldots, N$ unchanged. The exchangeable σ-algebra is $\mathcal{E} = \bigcap_{N \in \mathbb{N}} \mathcal{E}_N$.

Like many results in probability theory de Finetti's representation theorem is the consequence of a result on almost sure convergence. To appeal first a bit on intuition we begin investigating $\mathcal{E}$ in the i.i.d. case. The result is from Kallenberg, [9]. It is known as the 0-1-law of Hewitt-Savage. In this setting the Y_n are general random variables.

Theorem 3.3.2. *Let $(Y_n)_{n\in\mathbb{N}^*}$ be a sequence of* ***i.i.d.*** *random variables. Let $\mathcal{E}$ be the exchangeable σ-algebra. Then $\mathbb{P}(\mathcal{E}) = \{0, 1\}$.*

Proof. Let $E \in \mathcal{E}$. Let $E_n = C_n(E) \times \mathcal{X}^{\mathbb{N}}$ where $C_n(E)$ is the projection of E to the first n coordinates. Since $E_n \in \mathcal{E}_n$ it is measurable. The sequence (E_n) is monotonic and thus converges to E. We write $\tilde{E}_n$ for the set $\mathcal{X}^n \times E_n \times \mathcal{X}^{\mathbb{N}}$. For sets A and B denote by $A\Delta B$ the symmetric difference $(A\backslash B) \cup (B\backslash A)$. By exchangeability

$$\mathbb{P}(\tilde{E}_n\Delta E) = \mathbb{P}(E_n\Delta E) \to 0$$

and thus

$$\mathbb{P}(E_n \cap \tilde{E}_n) \to \mathbb{P}(E).$$

Observe that by joint independence of $(Y_n)_{n\in\mathbb{N}^*}$, $\mathbb{P}(E_n \cap \tilde{E}_n) = \mathbb{P}(E_n)\mathbb{P}(\tilde{E}_n)$. We have that

$$\mathbb{P}((Y_n) \in E) \leftarrow \mathbb{P}((Y_n) \in E_n \cap \tilde{E}_n) = \mathbb{P}((Y_n) \in E_n)\mathbb{P}((Y_n) \in \tilde{E}_n) \to \mathbb{P}^2((Y_n) \in E).$$

Hence $\mathbb{P}(E) = \{0, 1\}$. □

In the last theorem $\mathcal{E}$ was defined for i.i.d random variables. Of course, the definition of $\mathcal{E}$ is not restricted to i.i.d. or exchangeable sequences; the concept easily translates to arbitrary sequences of random variables defined on the same space. Observe that any sequence of i.i.d. random variables is trivially exchangeable. The converse does not hold true in general. Without assuming that the variables are i.i.d., the statement of the last lemma breaks down. But observe the similarity between the last theorem and Kolmogorov's 0-1-law for the shift-invariant σ-algebra. Suppose that we are already aware of the ocurrence of an event E in $\mathcal{E}$ and move the second element Y_2 to a different place N far ahead, leaving the order of the other Y's unchanged. By this we may hold Y_2 back further and further. So it is only natural to suspect that deleting Y_2 has no impact on E. By doing so we do not care much about Y_2, E is still true if we leave the rest of the Y's unchanged. Hence knowing that E occurs we might as well assume that Y_1 is independent of Y_2 since Y_2 was deleted anyway. The above reasoning leads to the following definition.

Definition 3.3.3. *Let $\mathcal{X}, \mathcal{B}(\mathcal{X})$ and (Y_n) be as in Defininition 3.2.1. Let $\mathcal{F} \subset \mathcal{B}(\mathcal{X})^{\mathbb{N}}$ be some sub-σ-algebra. We say that (Y_n) are* ***conditionally*** *i.i.d. if for all*

$N \in \mathbb{N}$, *for all* $A_i \in \mathcal{B}(\mathcal{X})$, $i = 1, \ldots, N$

$$\mathbb{P}\left(\bigcap_{i=1}^{n}[Y_i \in A_i]|\mathcal{F}\right) = \prod_{i=1}^{N} \mathbb{P}(Y_1 \in A_i|\mathcal{F}).$$

Conditional independence is the key property linking the two notions of exchangeability and mixture. The following two theorems have first been proved by de Finetti for sequences of Bernoulli variables. Ryll-Nardzewski proved it for general Polish space random variables, see [9]. A version for locally compact Hausdorff spaces was stated and proved by Hewitt and Savage, [8]. We apply the theorems with the sequence $(Y_n)_{n\in\mathbb{N}^*}$ from the beginning of the section. The notion of backwards martingales turns out to be useful. The approach of Hewitt and Savage is based on functional analysis while Ryll-Nardzewski's requires a rigorous proof of the reasoning above.

Theorem 3.3.4 (De Finetti, Ryll-Nardzewski)**.** *Let* $X = LRRW(G, a)$ *for some locally finite graph* G *and positive initial weights* a*. Assume that* X *is recurrent, so* $X = (Y_n)_{n\in\mathbb{N}}$*, where the* Y_n *are* x_0*-*x_0*-paths. Then* $(Y_n)_{n\in\mathbb{N}}$ *are conditionally independent given the exchangeable* σ*-algebra* $\mathcal{E}$*.*

Theorem 3.3.5. *Let the assumptions of the last theorem hold. Then* $(Y_n)_{n\in\mathbb{N}^*}$ *is a mixture of i.i.d. sequences.*

Definition 3.3.6. *A backwards filtration is a sequence of* σ*-algebras*

$$\ldots, \mathcal{F}_{-n} \subset \cdots \subset \mathcal{F}_{-1} \subset \mathcal{F}_0. \tag{3.9}$$

Let $M = (\ldots, M_{-2}, M_{-1}, M_0)$ *be a sequence of integrable, adapted, real-valued random variables. We call* M *a backwards martingale if*

$$\mathbb{E}(M_{-n}|\mathcal{F}_{-n-1}) = M_{-n-1} \tag{3.10}$$

holds almost surely.

There is a version of the martingale convergence theorem for backwards martingales. The use of negative indexing in (3.9) and (3.10) is little more than a formality. We will not give the whole proof but only the outlines. A version may be found in Rogers and Williams, [18].

Theorem 3.3.7. *Let* $(M_{-n})_{n\in\mathbb{N}}$ *be a uniformly integrable backwards martingale. Then* M_{-n} *converges almost surely to a random limit.*

Proof. Clearly, $E[M_0|\mathcal{F}_{-n}] = M_{-n}$ by the tower property of martingales, we may see $(M_k)_{k=-n,\dots,0}$ as a martingale. For $a < b \in \mathbb{Q}$ denote by $U_n(a,b)$ the number of upcrossings of the interval $[a,b]$ by $(M_{-n}, M_{-n+1}, \dots, M_0)$.

$$\mathbb{E}[U_n(a,b)] \leqslant \frac{\mathbb{E}[|M_0|] + a}{b-a}$$

holds by Doob's upcrossing inequality. Hence the number of upcrossings stays almost surely finite. Thus M_{-n} converges almost surely to a random limit $M_{-\infty}$. □

We are going to apply the theorem only using indicator functions on the space of finite paths. The following proof is due to Diaconis, [6, p.26–28].

Proof of Theorem 3.3.4. Denote by Y_n the projection of $(Y_k)_{k\in\mathbb{N}}$ to the n-th coordinate. Let $A \in \mathcal{B}(X)$. Define

$$M_n^A = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{Y_i \in A}.$$

For any set $A \in \mathcal{B}(Y)$ exchangeability yields

$$M_n^A = \mathbb{E}[\mathbb{1}_{Y_1 \in A}|\mathcal{E}_n]. \tag{3.11}$$

By substituting $\tilde{M}_{-n}^A := M_n(A)$ we note that $(\tilde{M}_n^A)$ is a backwards martingale. Since indicator functions are bounded this backwards martingale is uniformly integrable. Thus the limit M_∞^A exists and is given by

$$M_\infty^A = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{Y_i \in A} = \mathbb{E}[\mathbb{1}_{Y_i \in A}|\mathcal{E}].$$

We now show conditional independence of the Y_k. For $A_1, \dots, A_k \in \mathcal{B}(\mathcal{X})$ we define

$$f : Y^k \to \{0,1\}, f(y_1, \dots, y_n) = \prod_{i=1}^{k} \mathbb{1}_{y_i \in A_i}.$$

Write $n^{(k)}$ for $n(n-1)\cdots(n-k+1)$ and define

$$M_{-n} = \frac{1}{n^{(k)}} \sum_{\rho \in \mathfrak{S}_n} f(Y_{\rho(1)}, \dots, \rho(n)).$$

Like before, $(M_{-n})_{n\in\mathbb{N}}$ is a uniformly integrable backwards martingale and thus

$$\begin{aligned}\mathbb{E}[f(Y_1, \dots, Y_k)|\mathcal{E}] &= \lim_{n\to\infty} \frac{1}{n^{(k)}} \sum_{\rho \in \mathfrak{S}_n} f(Y_{\rho(1)}, \dots, Y_{\rho(n)}) = \\ &= \lim_{n\to\infty} \frac{1}{n^k} \sum_{\rho \in \mathfrak{S}_n} f(Y_{\rho(1)}, \dots, Y_{\rho(n)})\end{aligned}$$

Since all appearing terms are non-negative and the function $\sum_{i=1}^n \mathbb{1}_{Y_i \in A}$ is of order $\mathcal{O}(n)$ we may exchange summation, limit and product to get

$$\mathbb{P}(Y_1 \in A_1, \ldots, Y_k \in A_k | \mathcal{E}) = \prod_{i=1}^k \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{Y_i \in A_i} = \prod_{i=1}^k \mathbb{P}(Y_i \in A_i | \mathcal{E})$$

which finishes the proof. □

The idea of the following proof of Theorem 3.3.5 by construction is based loosely on Kallenberg, [9], and Hewitt and Savage, [8].

Proof of Theorem 3.3.5. We construct the measure space $(\mathcal{P}, \mathcal{B}(\mathcal{P}))$ via $\mathbb{P}$. For $\lambda \in [0,1] \cap \mathbb{Q}$ the event

$$M_A^\lambda = \left[M_\infty^A \leqslant \lambda\right] \tag{3.12}$$

is $\mathcal{E}$-measurable as well as countable unions and intersections of events of the form in (3.12). Let $\mathcal{P}$ be again the set of distributions on $\mathcal{B}(\mathcal{X})$ and define $\mathcal{B}(\mathcal{P})$ as the smallest σ-algebra containing all sets of the form

$$\mathcal{P}_{A,\lambda} := \{\pi \in \mathcal{P} : \pi(A) \in [0, \lambda]\}.$$

for all $A \in \mathcal{B}(\mathcal{X})$, for all $\lambda \in [0,1] \cap \mathbb{Q}$. For $(A_i)_{i \in \mathbb{N}} \subset \mathcal{B}(\mathcal{X})$ and $(\lambda_j)_{j \in \mathbb{N}} \subset [0,1]$ define for $\mathcal{P}' = \bigcup_{i=1}^\infty \bigcap_{j=1}^\infty \mathcal{P}_{A_i,\lambda_j}$

$$\mu(\mathcal{P}') := \mathbb{P}\left(\bigcup_{i=1}^\infty \bigcap_{j=1}^\infty M_{A_i}^{\lambda_j}\right).$$

Clearly, μ is a probability measure. Let $B = B_1 \times \cdots \times B_k \times \mathcal{X}^{\mathbb{N}}$ where $B_i \in \mathcal{B}(\mathcal{X})$ for $i = 1, \ldots, k$. Now, by Theorem 3.3.4, for any B of the form above

$$\mathbb{P}(B|E) = \prod_{i=1}^k \mathbb{E}(\mathbb{1}_{Y_i \in B_i} | E).$$

holds almost surely. If we choose $E \neq \varnothing$ of the form

$$E_\epsilon := \bigcap_{i=1}^k [M_\infty^{B_i} \in [p_i - \epsilon_i, p_i + \epsilon_i]]$$

for $p_i, \epsilon_i \in (0,1)$, $i = 1, \ldots, k$ then, for the above

$$\prod_{i=1}^k (p_i - \epsilon_i) \mathbb{P}(E_\epsilon) \leqslant \prod_{i=1}^n \mathbb{P}(B|E_\epsilon) \leqslant \prod_{i=1}^k (p_i + \epsilon_i) \mathbb{P}(E_\epsilon).$$

Now let P_B^ϵ be the set of probability distributions π on $\mathcal{B}(\mathcal{X})$ for which $p_i - \epsilon_i \leqslant \pi(B_i) \leqslant p_i + \epsilon_i, i = 1, \ldots, k$. With this notation the last equation simplifies to

$$\left(\prod_{i=1}^{k}(\pi(B_i) - \epsilon_i)\right)\mu(P_B^\epsilon) \leqslant \mathbb{P}(B|E) \leqslant \left(\prod_{i=1}^{k}(\pi(B_i) + \epsilon_i)\right)\mu(P_B^\epsilon).$$

Summing up for $\biguplus_{i=1}^{\infty} E_{i,\epsilon} = \mathcal{X}^{\mathbb{N}}$ and letting $\epsilon \to 0$ leads to

$$\mathbb{P}(B) = \int_{\mathcal{P}}\prod_{i=1}^{k}\pi(B_i)d\mu(\pi) = \int_{\mathcal{P}}\pi^{\mathbb{N}}(B)d\mu(\pi). \tag{3.13}$$

Since all $B \in \mathcal{B}(\mathcal{X})^{\mathbb{N}}$ may be approximated by finite unions of finite intersections of sets in $\mathcal{B}(\mathcal{X})^k \times \mathcal{X}^{\mathbb{N}}$ for some $k \in \mathbb{N}$ the existence of the measure μ with the disired properties is proved. We show almost sure uniqueness of μ on a generator. By (3.13)

$$\mathbb{P}(M_A^\lambda) = \int_{\mathcal{P}}\pi(M_A^\lambda)d\mu(\pi) = \int_{\mathcal{P}}\mathbb{1}_{\pi(A)\in[0,\lambda]}d\mu(\pi) = \mu(\mathcal{P}_{A,\lambda}),$$

where the equality in the middle follows by the strong law of large numbers. Since the sets $\mathcal{P}_{A,\lambda}$ generate $\mathcal{B}(\mathcal{P})$, μ is unique. $\square$

Now that it is proved that sequences of exchangeable random variables are mixtures of i.i.d. sequences we may apply this to show Theorem 3.1.1. The sequence of chosen balls is exchangeable and thus a mixture of i.i.d. sequences. Alternatively the relative content (i.e. the probability of choosing a specific ball) converges almost surely to a random limit by the Martingale Convergence Theorem. Hence to prove Theorem 3.1.1 it suffices to show that the density of the mixing measure coincides with the one of a Dirichlet distribution. The following proof has been obtained in the course of the elaboration of this thesis.

Proof of Theorem 3.1.1. We will prove the theorem by induction on m. The base is $m = 2$, let us just have red and white balls. For simplicity of notation let U_r be the amount of red balls and u_r and u_w the initial amounts of red and white balls, respectively. Denote by p the realization of the limit of $\bar{U}_{r,n}$, the relative amount of red balls. Since the sequence of chosen balls is exchangeable, we may write the probability of an event in terms of i.i.d. random variables and the mixing measure. The probability of an event $U_{r,n} = u_r + n_r$ in terms of mixtures (denoting by μ the mixing measure) is

$$\mathbb{P}(U_{r,n} = u_r + n_r) = \binom{n}{n_r}\int_0^1 p^{n_r}(1-p)^{n-n_r}d\mu(p). \tag{3.14}$$

Elementary combinatorial probability leads to

$$\mathbb{P}(U_{r,n} = u_r + n_r) = \binom{n}{n_r} \frac{\prod_{j=0}^{n_r-1}(u_r+j)\prod_{j=0}^{n-n_r-1}(u_w+j)}{\prod_{j=0}^{n-1}(u_w+u_r+j)} = \tag{3.15}$$

$$= \binom{n}{n_r} \frac{\beta(u_r+n_r, u_w+n-n_r)}{\beta(u_r,u_w)} \tag{3.16}$$

with β being the Beta-function defined by

$$\beta(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

Knowing that (3.14) and (3.15) are equal we may write

$$\int_0^1 p^{n_r}(1-p)^{n-n_r} d\mu(p) = \frac{\beta(u_r+n_r, u_w+n-n_r)}{\beta(u_r,u_w)} =$$
$$= \frac{1}{\beta(u_r,u_w)} \int_0^1 p^{u_r+n_r-1}(1-p)^{u_w+n-n_r-1} dp =$$
$$\int_0^1 p^{n_r}(1-p)^{n-n_r} \frac{p^{u_r-1}(1-p)^{u_w-1}}{\beta(u_r,u_w)} dp$$

Hence we obtain $f(p) = \frac{\Gamma(u_r+u_w)}{\Gamma(u_r)\Gamma(u_w)} p^{u_r-1}(1-p)^{u_w-1}$ or for $p = (p_r, p_w)$ essentially $f(p_r,p_w) = \frac{\Gamma(u_r+u_w)}{\Gamma(u_r)\Gamma(u_w)} p_r^{u_r-1}(p_w)^{u_w-1}$. Note that writing $f(p_r,p_w)$ instead of $f(p)$ does not change anything since $p_r + p_w = 1$. Now that the theorem is proved for $m = 2$ let us make the induction step.

$m \to m+1$
Suppose that we are suffering from red-green-blindness, the first colour is red and the second green, i.e. $p_1 = p_r + p_g$ and $u_1 = u_r + u_g$. Thus we have

$$f_U(p_r+p_g, p_2, \ldots, p_{m-1}, p_m) =$$
$$\frac{\Gamma(\sum_{i=1}^m u_i)}{\prod_{i=1}^m \Gamma(u_i)} (p_r+p_g)^{u_r+u_g-1} \prod_{i=2}^{m-1} p_i^{u_i-1}(1-p_r-p_g-p_2-\cdots-p_{m-1})^{u_m}. \tag{3.17}$$

By omitting all but red and green balls we obtain the same process as for two colours. Thus the probability of $[\bar{U}_r \leqslant z | \bar{U}_r + \bar{U}_g = p_1]$ is

$$\int_0^{\frac{p_r}{p_1}} \frac{s^{u_r-1}(1-s)^{u_g-1}}{\beta(u_r,u_g)} ds = \int_0^{p_r} \frac{\left(\frac{t}{p_1}\right)^{r-1}\left(1-\frac{t}{p_1}\right)^{g-1}}{\beta(u_r,u_g)} \frac{dt}{x_1} =$$
$$= \frac{1}{\beta(u_r,u_g)p_1^{u_r+u_g-1}} \int_0^{p_r} t^{u_r-1}(p_1-t)^{u_g-1} dt$$

Differentiating w.r.t. p_r and multiplying with (3.17) yields

$$f_U(p_r, p_g, p_2 \dots, p_m) =$$
$$= \frac{\Gamma(u_r + u_g + \sum_{i=2}^m u_i)}{\Gamma(u_r)\Gamma(u_g)\prod_{i=2}^m \Gamma(u_i)} p_r^{u_r-1} p_g^{u_g-1} \prod_{i=2}^{m-1} p_i^{u_i-1} (1 - p_r - p_g - p_2 - \dots - p_{m-1})^{u_m-1}$$

Renaming

$$u_1' = u_r, u_2' = u_g, u_3' = u_2, \dots, u_{m+1}' := u_m,$$
$$p_1' := p_r, p_2' := p_g, p_3' := p_2, \dots, p_{m+1}' := p_m$$

finishes the proof. □

3.4 Diaconis, Freedman and the Recurrent Case

In the last section we obtained a representation of exchangeable sequences of random variables as a mixture of i.i.d. random variables. To generalizes the concept we consider these i.i.d. random variables to be paths on a graph. For the proof strategy the condition recurrence is crucial. The first part of the proof is due to Diaconis, [6], the part about reversibility due to Rolles, [19].

Theorem 3.4.1. *Let $X = LRRW(G, a)$ be recurrent. Then X is a unique mixture of reversible Markov chains.*

Proof. Recall the meaning of $\mathcal{X}$, $\mathcal{B}(\mathcal{X})$ and Y from the beginning of the last section. By recurrence, X is almost surely a sequence of $\mathcal{X}$-valued random variables. Consider the process $(Y_1, Y_2, \dots)$ which arises naturally by cutting X into pieces, each piece contained in $\mathcal{X}$. This sequence is a mixture of i.i.d. random variables. For the rest of the proof we assume that such a sequence of i.i.d. random variables with values in $\mathcal{X}$ is given and denote their probability distribution by P. We need to show that this sequence corresponds to a reversible Markov chain. Let y_1, y_2 be finite paths that end in the same state z and let $x_j \in \mathcal{N}_G(z)$. Although y_1 and y_2 are not contained in $\mathcal{X}$ they correspond to a set of paths in $\mathcal{X}$ and thus may be seen as P-measurable events. We show the Markov property

$$P([y_1, x_j] | y_1) = P([y_2, x_j] | y_2)$$

or, avoiding division by 0,

$$P(y_1)P([y_2, x_j]) = P(y_2)P([y_1, x_j]).$$

Let $\mathcal{X}^z$ be the set of finite paths starting in z, ending in x_0 and not visiting x_0 in between. Since X is recurrent

$$P(y_1) = \sum_{y \in \mathcal{X}^z} P([y_1, y]).$$

Observe that for $y', y'' \in Y$, $P([y', y'']) = P(y')P(y'')$ and hence for $y \in \mathcal{X}^z$

$$\begin{aligned} P\left([y_1, y]\right) P([y_2, x_j]) &= P([y_1, y, y_2, x_j]) = \\ &= P([y_2, y, y_1, x_j]) = P([y_2, y])P([y_1, x_j]). \end{aligned}$$

Summing up for $y \in \mathcal{X}^z$ we get

$$\begin{aligned} P(y_1)P([y_2, x_j]) &= \sum_{y \in \mathcal{X}^z} P([y_1, y])P([y_2, x_j]) \\ &= \sum_{y \in \mathcal{X}^z} P([y_2, y])P([y_1, x_j]) = P(y_2)P([y_1, x_j]). \end{aligned}$$

Since the lengths of y_1 and y_2 are arbitrary, also time homogeinity holds.
To prove reversibility we will make use of cycles c in Y. Note that by partial exchangeability the probability of any cycle starting and ending in x_0 does not depend on the orientation. Let c be a cycle and c^- be its reversal. By recurrence one will arrive upon an element of c. Denote by $Q_P(c)$ the probability w.r.t. the Markov chain with transition matrix P that the process performs c. We denote by μ the mixing measure. By partial exchangeability

$$q := \int_{\mathcal{P}} Q_P^2(c) d\mu(P) = \int_{\mathcal{P}} Q_P(c) Q_P(c^-) d\mu(P) = \int_{\mathcal{P}} Q_P^2(c^-) d\mu(P).$$

Thus we may write

$$\int_{\mathcal{P}} (Q_P(c) - Q_P(c^-))^2 d\mu(P) = q - 2q + q = 0.$$

Since the integrand is non-negative we conclude that $Q_P(c) = Q_P(c^-)$ almost surely for all cycles c. We still need to show that μ is a Borel-measure. We refer to $\mathcal{P}_{v,w}$ as the set reversible stochastic matrices $P = p(\cdot,\cdot)$ for which $p(v, w) \in B$ for some Borel set $B \subset [0, 1]$. $\mathcal{P}_{v,w}$ is exactly the set of distributions π on $\mathcal{X}$ for which

$$\pi(w \text{ is visited after the first visit to v}) \in B$$

since the event $[v$ is visited$]$ is an almost sure one. Sets of this form generate $\mathcal{B}(\mathcal{P})$. □

Remark 3.4.2. *If we choose the definition of partial exchangeabiliy in the sense of Diaconis and Freedman everything works fine up to reversibility.*

3.5 The Set of Reversible Markov Chains

There is no need to represent a reversible Markov chain by its associated transition matrix. We may as well define a **reversible** chain if we choose the weight of one edge e_1 to be 1 and arbitrary positive weights $\mathbb{W} = (\mathbb{W}_e)_{e \in E(G)}$ for all other edges. By setting

$$p(u,v) = \frac{\mathbb{W}_{\{u,v\}}}{\sum_{w \in \mathcal{N}_G(u)} \mathbb{W}_{\{u,w\}}} \tag{3.18}$$

the resulting matrix is the transition matrix of a reversible Markov chain. Let G be a locally finite connected undirected graph. Since the notation is easier in terms of directed graphs we substitute each edge in G by a pair of oppositely directed edges. We will denote by $\delta_G^+(v)$ the set of edges going out of a vertex v. $P = p(\cdot,\cdot)$ is the transition matrix of a reversible chain if and only if it is a solution to the system

$$\begin{aligned} \forall v \in V(G) : &\quad \sum_{e \in \delta_G^+(v)} p(e) = 1 \\ \forall \text{ cycles } c \in G : &\quad \prod_{e \in c^+} p(e) = \prod_{e \in c^-} p(e) \\ \forall e \in E(G) : &\quad p(e) \geqslant 0. \end{aligned}$$

Indeed, the system is quite easy to solve for trees since no cycles exist. Choosing an arbitrary set P_v of $\deg_G(v) - 1$ transition probabilities $p(v,\cdot)$ for each vertex $v \in V(G)$ that fulfill $\sum_{p \in P_v} p \leqslant 1$, this determines the other transition probabilities. But already in case G is a triangle the situation gets much more complicated. For graphs containing cycles the set of non-linear equations in the second line become cumbersome. However, irrespective of the last paragraph, μ is a mixing measure on $\mathcal{B}(W)$, where $W = (0,\infty)^{E(G)\setminus\{e_1\}}$ and $\mathcal{B}(W)$ is the σ-algebra generated by sets of the form

$$\{w = (w_e)_{e \in E(G)} | w_{e_0} \in A \subset (0,\infty), A \text{ open}, e_0 \in E(G)\}. \tag{3.19}$$

for $e_0 \in E(G)$. By the above isomorphism we may always choose the more convenient notation.

3.6 Pólya, Merkl, Rolles and the General Case

In case of trees $T = (V,E)$ the existence of a representation in the general case is almost trivial. The measure μ^* induced by the product of countably many Dirichlet distributions $\prod_{v \in V(T)} D_v$ fulfills the properties of a mixing measure since for any finite path we could as well assume that G is finite. In this case the reasoning at the beginning of Section 3.1 and Theorem 3.1.1 imply that μ^* is a mixing measure. The existence of a representation in terms of Markov chains without the assumption of recurrence is non-trivial in case G contains cycles. These generate dependencies between the edge weights. But providing sufficient conditions for recurrence or transience is much easier given a representation, so for the latter constructing a mixing measure is highly useful, motivating the statement of the following theorem.

Theorem 3.6.1. *On any locally finite graph G, for all initial weights $0 < a = (a_e)_{e\in E(G)}$ $LRRW(G, a)$ is a mixture of irreducible reversible Markov chains, no matter if it is recurrent, transient or maybe neither.*

Apart from the following example the results in this section, including the statement above, are due to Merkl and Rolles, [15]. We give a short summary of the proof strategy. First we compare the distribution of transition probabilities from a specific vertex v for $LRRW$ with the distribution of transition probabilities in case of a star graph. For this comparison we introduce a concept called convex order. We then consider mixing measures $\mu^{(n)}$ on an increasing sequence of subgraphs G_n of G. We will argue later that on finite graphs $LRRW$ is recurrent and thus these measures exist. A tightness result leads to the construction of a weak limit μ^* of a subsequence. We prove that μ^* fulfills the desired properties, finishing the chapter proving that $LRRW$ is recurrent if and only if it contains a recurrent state. We use the notation introduced in the last section. μ^* will be a measure on $(0, \infty)^{E(G)\setminus\{e_1\}}$.
But first, let us provide a counterexample of a partially exchangeable process that is not a mixture of Markov chains. The example is due to Diaconis and Freedman, [6], and shows that Theorem 3.6.1 breaks down for general partially exchangeable processes.

Example 3.6.2. *Let $G = (\{0, 1\}, \{\{0, 1\}, \{0\}, \{1\}\})$ be the graph consisting of one edge and two loops. Observe first the following transient process X that starts at 0 and repeats 0 finitely often before switching to 1 and staying there forever. Denote the probability that X walks along the transient path starting with k zeroes before switching to 1 by $m_k(X)$. Each path is uniquely determined by its initial vertex and the transition numbers, hence partial exchangeability holds trivially. If X is a mixture then the mixing measure μ puts strictly positive mass δ on $p(1, 1) = 1$. Therefore m_k can be written as*

$$m_k = \int_{[p(1,1)=1]} (p(0, 0))^{k-1} p(0, 1) d\mu(p) = \delta \int_0^1 (p(0, 0))^{k-1} (1 - p(0, 0)) d\mu(p(0, 0)).$$

Observe that m_k is decreasing in k. Now consider the degenerate process $X' = (0,0,1,1,1,1,1,\dots)$. We show that X' is not a mixture of Markov chains. Trivially, X' is partially exchangeable, so $m_k(X')$ is decreasing in k. But $m_1(X') = 0$, $m_2(X') = 1$ and $m_3(X') = 0$. Contradiction to monotonicity. However, X' is not LRRW, it does not even meet the assumptions of any reinforced random walk.

3.6.1 Convex Order and Pólya's Urns again

This section will deal a lot with the routines on the edges. For simplicity we denote by

$$w_v(t) = \sum_{e \in \delta_G(v)} w_e(t) \tag{3.20}$$

the routine of edge e at time t and the sum of routines of edges incident to v, respectively. The following result has been elaborated in the course of this thesis.

Lemma 3.6.3. *On any **finite** connected graph $G = (V,E)$, $X = LRRW(G,a)$ is recurrent for all strictly positive $(a_e)_{e \in E}$.*

Proof. For an infinite path y let $V_\infty := V_\infty(y)$ be the set of vertices are visited by y infinitely often. Clearly, $V_\infty \neq \varnothing$ holds always true. We show that $\mathbb{P}(V_\infty \neq V(G)) = 0$. Let $v, v_0 \in V(G)$, $\{v, v_0\} \in E(G)$. Let $t \in \mathbb{N}$. We write $W_v(t)$ and $W_{v,v_0}(t)$ for $w_v(t)$ and $w_{v,v_0}(t)$, considered random variabels. We have that

$$\mathbb{P}(X_{t'} \neq v_0 \text{ for all } t' > t | v \in V_\infty, W_v(t), W_{v,v_0}(t)) \leqslant \tag{3.21}$$

$$\leqslant \prod_{i=0}^{\infty} \left(1 - \frac{W_{v,v_0}(t)}{W_v(t) + 2i - 1}\right) \leqslant \exp\left(-\sum_{i=0}^{\infty} \frac{W_{v,v_0}(t)}{W_v(t) + 2i - 1}\right) = 0. \tag{3.22}$$

Observe that (3.21) holds true for all possible values of $W_v(t)$ and $W_{v,v_0}(t)$. Hence

$$\mathbb{P}(\exists T \in \mathbb{N} : X_t \neq v_0 \text{ for all } t > T | v \in V_\infty) = 0.$$

Thus, if $v \in V_\infty$ then $v' \in V_\infty$ holds almost surely for all $v' \in \mathcal{N}_G(v)$ and, iteratively, for all $v' \in V(G)$. Observing again that $V_\infty \neq \varnothing$ finishes the proof. □

Consider $X = LRRW(G,a)$ for initial vertex x_0 on a **finite** graph G. Let $S = G[\{v\} \cup \mathcal{N}_G(v)]$ be the induced star graph with center v. Reconsider Pólya's urn model from Section 3.1. Note that if G is a tree the routines of the edges in S behave essentially like a Pólya urn model with increment 2 instead of 1. Since the mixing measure in this case has a closed form a comparison between these two models is desirable.

Remark 3.6.4. *In the last paragraph assuming $v = x_0$ is not very restrictive. If it does not hold let e^* be the edge by which v is reached first and t^* the first time v is visited. Since G is finite $t^* < \infty$ almost surely by the statement of the last lemma. Take*

$$X'_n = X_{n+t^*}, \quad a'_e = \begin{cases} a_e + 1 & \text{if } e = e^* \\ a_e & \text{else} \end{cases} \tag{3.23}$$

and realize X' on (G, a'). Troughout this chapter, each time we talk about an (induced) star graph we will assume without further mentioning that the process starts in the center.

Fix a vertex v and an edge e incident to v. Let T_i be the $i+1$-st visit to v. We define

$$M_n := \frac{w_e(T_n)}{w_v(T_n)} \tag{3.24}$$

Now consider the star graph $S = G[\{v\} \cup \mathcal{N}_G(v)]$ induced by v and its adjacent vertices. Let $X^S = LRRW(S, (a_e)_{e \in E(S)})$ with $x_0^S = v$. Recall remark 3.6.4. Define $w_e^S(t)$, $w_v^S(t)$ and $M_n^S(t)$ analogously to (3.20) and (3.24), i.e.

$$w_e^S(t) = a_e + \sum_{i=1}^{t} \mathbb{1}_e(\{X_{i-1}^S, X_i^S\}) \qquad\qquad w_v^S(t) = \sum_{e \in E(G^S)} w_e(t)$$
$$M_n^S(t) = \frac{w_e^S(T_n^S)}{w_v^S(T_n^S)}.$$

Define filtrations $(\mathcal{G}_n)_{n \in \mathbb{N}}$ and $(\mathcal{G}_n^S)_{n \in \mathbb{N}}$ by

$$\mathcal{G}_n = \sigma(w_e(T_k) : k = 0, \ldots, n) \text{ and}$$
$$\mathcal{G}_n^S = \sigma(w_e^S(T_k) : k = 0, \ldots, n).$$

Lemma 3.6.5. *M_n^S is a martingale w.r.t. $(\mathcal{G}_n^S)$ and M_n is a martingale w.r.t. $(\mathcal{G}_n)$.*

Proof. The second statement is a special case of the first, hence showing that M_n is a martingale suffices. Let $\mathcal{Y}$ be the set of paths that return to v infinitely often. Consider the following partition of $\mathcal{Y}$.

- $A_n = [\{X_{T_n}, X_{T_n+1}\} = e, \{X_{T_{n+1}-1}, X_{T_{n+1}}\} = e]$
- $B_n = [\{X_{T_n}, X_{T_n+1}\} = e, \{X_{T_{n+1}-1}, X_{T_{n+1}}\} \neq e]$
- $C_n = [\{X_{T_n}, X_{T_n+1}\} \neq e, \{X_{T_{n+1}-1}, X_{T_{n+1}}\} = e]$
- $D_n = [\{X_{T_n}, X_{T_n+1}\} \neq e, \{X_{T_{n+1}-1}, X_{T_{n+1}}\} \neq e]$

Denote by $(\mathcal{F}_n)_{n\in\mathbb{N}}$, $\mathcal{F}_n \supset \mathcal{G}_n$, some filtration so that the elements of the partition are measureable w.r.t. $\mathcal{F}_n$. Each element of the partition corresponds to a set of finite paths, hence $\mathcal{F}_n$ is exists. Note that the transformation τ_n that reverts the last (v,v)-path before T_{n+1} is an isomorphism on $\mathcal{Y}$. A_n and D_n are mapped onto themselves respectively, B_n is mapped to C_n bijectively. By partial exchangeability τ_n is measure preserving, hence $\mathbb{P}(B|\mathcal{F}_n) = \mathbb{P}(C|\mathcal{F}_n)$. Observe as well that by starting from v, $\mathbb{P}(A_n \cup B_n|\mathcal{F}_n) = M_n$. So

$$\begin{aligned}2\mathbb{P}(A_n|\mathcal{F}_n) + \mathbb{P}(B_n \cup C_n|\mathcal{F}_n) =&\mathbb{P}(A_n \cup B_n|\mathcal{F}_n) + \mathbb{P}(A_n \cup C_n|\mathcal{F}_n) = \\ =&2\mathbb{P}(A_n \cup B_n|\mathcal{F}_n) = 2M_n,\end{aligned}$$

and hence

$$\begin{aligned}\mathbb{E}[M_{n+1}|\mathcal{F}_n] = &\frac{w_e(T_n) + 2P(A_n|\mathcal{F}_n) + P(B_n \cup C_n|\mathcal{F}_n)}{w_v(T_{n+1})} = \\ =&\frac{M_n(w_v(T_n)+2)}{w_v(T_{n+1})} = M_n.\end{aligned}$$

By definition M_n is measureable w.r.t. $\mathcal{G}_n$, so writing $\mathcal{G}_n$ in the first and the last term instead of $\mathcal{F}_n$ does not change anything. □

Let us first give an intuitive comparison of M_n and M_n^S. Assume that at each time T_n we could treat M_n as if it was a result of X^S. This means, that by coincidence X left v the same number of times via e as it returned via e, i.e. $w_e(T_n) = w_e^S(T_n)$, hence $M_n = M_n^S$. For a specific v-v-path c entering and leaving v via different edges we denote by c^+ and c^- directed versions and by $N(c^+, n)$ and $N(c^-, n)$ the number of walks along c^+ and c^-, respectively. We note again that *LRRW* is a mixture of reversible Markov chains. For a fixed reversible chain the ratio

$$\frac{N(c^+,n)}{N(c^+,n)+N(c^-,n)}$$

converges almost surely to $\frac{1}{2}$ by the law of large numbers since c^+ and c^- have the same probability. However, the M_n and M_n^S are not defined on the same probability space. Apart from that M_n^S is much coarser than M_n. This observation motivates the definition of convex order.

Definition 3.6.6 (Convex Order)**.** *Let U, V be integrable random variables, not necessarily on the same probability space. Denote by $\overset{d}{=}$ equality in distribution. We say that U and V are in convex order, $U \lhd V$, if there are random variables $U_1 \overset{d}{=} U$ and $V_1 \overset{d}{=} V$ on the same probability space $(\Omega, \mathcal{A}, P)$ so that (U_1, V_1) is a 1-step-martingale. If μ and ν are the probability measures associated with U and V, respectively, we also write $\mu \lhd \nu$.*

Theorem 3.6.7. *Denote the random limits of M_n and M_n^S by M and M^S, respectively. These exist by the martingale convergence theorem.*

$$M_n \lhd M_n^S$$

holds true for all $n \in \mathbb{N}$ and

$$M \lhd M^S.$$

Obviously $M_0 \lhd M_0^S$ since they are equal. M_n and M_n^S are not defined on the same space. However, the statement implies that there is a space on which it is possible to define M_n and M_n^S so they fulfill the properties of $LRRW$ and Pólya's urn model, respectively. Therefore we made an attempt to find a constructive proof of Theorem 3.6.7. The result was anything but satisfactory. Instead, it turns out to be convenient to introduce the Pólya urn transition kernel

$$K_n : [0,1] \times \mathcal{B}([0,1]) \to [0,1], \qquad K_n(x,\cdot) = x\mathbb{1}_{\lambda_n x+1-\lambda_n} + (1-x)\mathbb{1}_{\lambda_n x}, \quad (3.25)$$

where $\lambda_n := \frac{a_v+2n}{a_v+2n+2}$. We are going to use K_n for both M_n^S and M_n. In the case of M_n^S the resulting process **is** in fact M_n^S whereas in case of M_n this results in a new random variable $\tilde{M}_n$, defined as follows.

$$\tilde{M}_0 := M_0, \quad \tilde{M}_{n+1} := \frac{w_e(T_n) + 2\cdot \mathbb{1}_{A_n \cup B_n}}{w_v(T_{n+1})} = \lambda_n M_n + \mathbb{1}_{A_n\cup B_n}(1-\lambda_n). \quad (3.26)$$

Given M_n, the value of $\tilde{M}_{n+1}$ depends only on the edge by which X exits v at time T_n. We would like to show that for all $n \in \mathbb{N}$

$$M_n \lhd \tilde{M}_n \text{ and } \tilde{M}_n \lhd M_n^S$$

and conclude that $M_n \lhd M_n^S$ and furthermore $M \lhd M^S$. To be able to do so we show 4 properties of $\lhd$ in the next lemma. It is clear that we need transitivity of $\lhd$. For the conclusion

$$(\forall n \in \mathbb{N} \quad M_n \lhd M_n^S) \Rightarrow (M \lhd M^S)$$

we need to show that $\lhd$ is preserved when we pass M_n and M_n^S to their respective limits. Proving $M_n \lhd \tilde{M}_n$ is not a difficult task and hence shown in the proof of Theorem 3.6.7. Part 4 requires Part 3 and is needed to show

$$M_n \lhd M_n^S \Rightarrow \tilde{M}_{n+1} \lhd M_{n+1}^S. \quad (3.27)$$

Lemma 3.6.8. *1. $\lhd$ is transitive.*

2. Let $(U_n)_{n\in\mathbb{N}}$ and $(V_n)_{n\in\mathbb{N}}$ be uniformly integrable martingales w.r.t. their natural filtrations $\mathcal{F}_n$ and $\mathcal{G}_n$, respectively. If $U_n \lhd V_n$ holds true for all $n \in \mathbb{N}$ then their respective limits U and V satisfy $U \lhd V$.

3. *For $i = 1, \dots, n$ let $\mu_i \lhd \nu_i$ be probability distributions on $\mathbb{R}$, let $p \in \mathbb{R}^n_+$, $||p||_1 = 1$ be a stochastic vector. Then $\sum_{i=1}^n p_i\mu_i \lhd \sum_{i=1}^n p_i\nu_i$.*

4. *Let μ, ν be probability measures on $([0,1], \mathcal{B}([0,1]))$, μ discrete on a finite set and $\mu \lhd \nu$. Let K_n be the Pólya transition kernel introduced in (3.25). Then $\mu K_n \lhd \nu K_n$.*

Proof. 1. Let $U \lhd V \lhd W$ be random variables and let (U_1, V_1) and (V_2, W_2) the corresponding 1-step-martingales. Denote by $P_1(V_1|U_1)$ the conditional probability of U_1 given V_1. Define $P_2(V_2|W_2)$ analogously. Now let P_3 be the law of a 3-point process $M = (U^*, V^*, W^*)$ given by

$$P_3(A \times B \times C) = \int_A \int_B P_2(W_2 \in C|V_2 = v)P_1(V_1 \in dv|U_1 = u)dP(u)$$

Denote by U_3, V_3, W_3 the projections onto the first, second and third coordinate, respectively. Observe that M is a Markov chain and $U_3 \stackrel{d}{=} U, V_3 \stackrel{d}{=} V, W_3 \stackrel{d}{=} W$ by construction. By the tower property for martingales

$$\mathbb{E}_{P_3}[W_3|U_3] = \mathbb{E}_{P_3}[\mathbb{E}_{P_3}[W_3|V_3, U_3]|U_3] = \mathbb{E}_{P_3}[V_3|U_3] = U_3$$

Thus $U \lhd W$.

2. Observe first that, since V_n is uniformly integrable, $E[V|\mathcal{G}_n] = V_n$ and thus $V_n \lhd V$. We have $U_n \lhd V$ because $U_n \lhd V_n \lhd V$. Thus there are $U'_n \stackrel{d}{=} U$ and $V'_n \stackrel{d}{=} V$ so that $U'_n = \mathbb{E}_{P_n}[V'_n|U'_n]$ where P_n denotes the measure on a common probability space. Equivalently, for any bounded continuous function $f : \mathbb{R} \to \mathbb{R}$

$$\mathbb{E}_{P_n}[f(U'_n)U'_n] = \mathbb{E}_{P_n}[f(U'_n)V'_n], \tag{3.28}$$

$(U'_n)_{n\in\mathbb{N}}$ is uniformly integrable and therefore a tight sequence, the same holds for V'_n. Thus there is a weakly convergent subsequence $(U'_{n_k}, V'_{n_k})_{k\in\mathbb{N}}$. Denote its limit by (U', V') and the probability measure by P. Taking the limit in (3.28) yields

$$\mathbb{E}_P[f(U')U'] = \mathbb{E}_P[f(U')V'],$$

again equivalent to $U' = \mathbb{E}[V'|U']$. Thus $U \lhd V$.

3. For $i = 1, \dots, n$ let $X_i \lhd Y_i$ be random variables on a probability space $(\Omega_i, \mathcal{A}_i, \mathbb{P}_i)$ with laws μ_i and ν_i, respectively. Since for all $i = 1, \dots, n$ $\mu_i \lhd \nu_i$ holds there exist a sub-σ-algebras $\mathcal{F}_i$ so that $X_i = \mathbb{E}[Y_i|\mathcal{F}_i]$. W.l.o.g. we

assume that $(X_i)_{i=1,\dots,n}$ and $(Y_i)_{i=1,\dots,n}$ are jointly independent, respectively, and that the Ω_i are pairwise disjoint. We take

$$\Omega := \bigcup_{i=1}^{n} \Omega_i, \qquad \mathcal{A} := \sigma\left(\bigcup_{i=1}^{n} \mathcal{A}_i\right)$$
$$\mathbb{P}(A) := p_i \mathbb{P}_i(A \cap \Omega_i), \qquad \mathcal{F} := \sigma\left(\bigcup_{i=1}^{n} \mathcal{F}_i\right).$$

Define $X = \sum_{i=1}^{n} X_i \mathbb{1}_{\Omega_i}$, $Y = \sum_{i=1}^{n} Y_i \mathbb{1}_{\Omega_i}$.
$\mu = \sum_{i=1}^{n} p_i \mu_i$, $\nu = \sum_{i=1}^{n} p_i \nu_i$ and thus $\mathbb{E}[Y|\mathcal{F}] = X$.

4. Let us first prove this for $\mu = \mathbb{1}_x$. We want to show that if ν has expectation x then
$$\mu K_n = x\mathbb{1}_{\lambda_n x + 1 - \lambda_n} + (1-x)\mathbb{1}_{\lambda_n x} \lhd \nu K_n. \tag{3.29}$$
For this purpose we introduce random variables X, Y with joint distribution $\nu \otimes K_n$. We observe that $Y < X$ is equivalent to $Y = \lambda_n X$ and $Y \geqslant X$ is equivalent to $Y = \lambda_n X + 1 - \lambda_n$. By this we deduce that
$$\begin{aligned}
\mathbb{P}(Y \geqslant X) &= \int_{[0,1]} y d\nu(y) = x \\
\mathbb{E}[Y, Y \geqslant X] &= \int_{[0,1]} (\lambda_n y + 1 - \lambda_n) y d\nu(y) = (\lambda_n x + 1 - \lambda_n)x + \lambda_n \mathrm{Var}[\nu] \\
\mathbb{P}(Y < X) &= 1 - x \\
\mathbb{E}[Y, Y < X] &= \int_{[0,1]} \lambda_n y(1-y) d\nu(y) = \lambda_n x(1-x) - \lambda_n \mathrm{Var}[\nu].
\end{aligned}$$
Thus we have shown that
$$x\mathbb{1}_{\lambda x + 1 - \lambda + \lambda\frac{\mathrm{Var}[\nu]}{x}} + (1-x)\mathbb{1}_{\lambda x - \lambda\frac{\mathrm{Var}[\nu]}{1-x}} \stackrel{d}{=} \mathbb{E}[Y|\sigma([Y \geqslant X])] \lhd Y \sim \nu K_n$$
By transitivity of $\lhd$ it suffices to show that
$$x\mathbb{1}_{\lambda x + 1 - \lambda} + (1-x)\mathbb{1}_{\lambda x} \lhd x\mathbb{1}_{\lambda x + 1 - \lambda + \lambda\frac{\mathrm{Var}[\nu]}{x}} + (1-x)\mathbb{1}_{\lambda x - \lambda\frac{\mathrm{Var}[\nu]}{1-x}} \tag{3.30}$$
in order to prove (3.29). Let, more generally, $A \leqslant B \leqslant C \leqslant D \in \mathbb{R}$ satisfying
$$xA + (1-x)D = xB + (1-x)C.$$
We show (3.30) by
$$x\mathbb{1}_B + (1-x)\mathbb{1}_C \lhd x\mathbb{1}_A + (1-x)\mathbb{1}_D. \tag{3.31}$$

For shortness we will write $\Delta_{xy} = x - y$. Let $\Omega = \{B, C\} \times \{A, D\}$, denote its elements by (ω_1, ω_2). Define a measure μ_Ω on the power set of Ω by

$$\mu_\Omega = \frac{x\Delta_{DB}}{\Delta_{DA}}\mathbb{1}_{BA} + \frac{x\Delta_{BA}}{\Delta_{DA}}\mathbb{1}_{BD} + \frac{(1-x)\Delta_{DC}}{\Delta_{DA}}\mathbb{1}_{CA} + \frac{(1-x)\Delta_{CA}}{\Delta_{DA}}\mathbb{1}_{CD}$$

Note that in case $A = D$, (3.31) is trivial. Otherwise

$$\mathbb{E}[\omega_2|\omega_1 = B] = B \text{ and } \mathbb{E}[\omega_2|\omega_1 = C] = C$$

and 3.30 is proved. Now instead of $\mu = \mathbb{1}_x$ we take $\mu = \sum_{i=1}^n p_i \mathbb{1}_{x_i}$. If $X \sim \mu$ and $Y \sim \nu$ with $X = \mathbb{E}[Y|X]$ define

$$\nu_i := P[Y \in \cdot|X = x_i]$$

Since the claim holds for one-point-measures,

$$\mathbb{1}_{x_i} K_n \lhd \nu_i K_n.$$

Since μ is discrete on a finite set,

$$\mu = \sum_{i=1}^n p_i \mathbb{1}_{x_i}$$

for some stochastic vector $p \in [0,1]^n$, $||p||_1 = 1$. Using Part 3 of the lemma,

$$\mu K_n = \sum_{i=0}^\infty p_i \mathbb{1}_{x_i} K_n \lhd \sum_{i=1}^n p_i \nu_i K_n = \nu K_n.$$

$\square$

Proof of Theorem 3.6.7. The proof is by induction. Like mentioned before it is obvious that $M_0 \lhd M_0^S$. Recall the partition $\{A_n, B_n, C_n, D_n\}$ and the isomorphism τ from Lemma 3.6.5. For the induction step consider again the random variable $\tilde{M}_n$ defined in (3.26). We would like to show that $M_{n+1} \lhd \tilde{M}_{n+1}$. Indeed, $M_{n+1}\mathbb{1}_{A_n \cup D_n} \equiv \tilde{M}_{n+1}\mathbb{1}_{A_n \cup D_n}$ and hence the claim holds trivially on A and on D, which are both measureable w.r.t. $\mathcal{G}_{n+1}$. But observe that even on $B \cup C$

$$\tilde{M}_{n+1} + \tilde{M}_{n+1} \circ \tau_{n+1} = 2 \cdot M_{n+1}.$$

Now observe that $\mathcal{G}_{n+1}$ does not distinguish between the cases B and C. We conclude

$$\mathbb{E}[\tilde{M}_{n+1}|\mathcal{G}_n] = \mathbb{E}[\frac{1}{2}(\tilde{M}_{n+1} + \tilde{M}_{n+1} \circ \tau_{n+1})|\mathcal{G}_{n+1}] = M_{n+1}.$$

We would like to show $\tilde{M}_{n+1} \lhd M^S_{n+1}$. Recall Part 4 of Lemma 3.6.8. By the induction hypothesis $M_n \lhd M^S_n$ we get

$$\tilde{M}_{n+1} = K_n(M_n, \cdot) \lhd K_n(M^S_n, \cdot) = M^S_{n+1}.$$

and by transitivity of $\lhd$, $M_{n+1} \lhd M^S_{n+1}$. Applying Part 2 of Lemma 3.6.8 yields

$$M \lhd M^S.$$

□

To make use of Theorem 3.6.7 we need to determine the distribution of M^S. Recall Remark 3.6.4 and recall that in the two models M_n and M^S_n the initial weights a need to be adjusted to a' according to the edge by which v is first entered in the case $v \neq x_0$. Then in order to compare M and M^S recall that M_S behaves like Pólya's urn with increment 2 instead of 1. Taking half the initial weight and increment $1/2$ instead of 1 leaves $LRRW$ unchanged. Hence M^S_n behaves like Pólya's urn with initial amount of balls $\frac{a'_e}{2}$ and $\frac{a'_v - a'_e}{2}$. Theorem 3.1.1 yields

$$M^S \sim \begin{cases} \beta(\frac{a_e}{2}, \frac{a_v - a_e}{2}) & \text{if } X_0 = v \\ \beta(\frac{a_e+1}{2}, \frac{a_v - a_e}{2}) & \text{if } X_0 \neq v, w_e(T^v_0) = a_e + 1 \\ \beta(\frac{a_e}{2}, \frac{a_v - a_e + 1}{2}) & \text{if } X_0 \neq v, w_e(T^v_0) = a_e \end{cases}. \tag{3.32}$$

Note that

$$\frac{W_e(t)}{W_v(t)} =: M^{v,e}_{\max n: T_n \leqslant t} \xrightarrow[t \to \infty]{} M^{v,e}$$

hold almost surely, where $M^{v,e}_n$ is the martingale in (3.24) defined via e and v. $M^{v,e}$ is the almost sure limit of the probability w.r.t. a Markov chain of leaving v via e. In the last chapter we made it clear that a reversible Markov chain may always be defined in terms of non-negative weights on the edges.

3.6.2 Tightness of the Mixing Measures

Lemma 3.6.9. *Let $v \in V(G)$ and $e \in \delta_G(v)$, let $a = (a_e)_{e \in E(G)}$ and $M^{v,e}$ as above. Then for $LRRW(G, a)$ there exist constants $\kappa_1(a_e, a_v), \kappa_2(a_e, a_v)$ continuous in both arguments so that for all $\varepsilon > 0$*

$$\mathbb{P}\left(M^{v,e} \leqslant \varepsilon\right) \leqslant \kappa_1(a_v, a_e)\varepsilon^{a_e/2} \text{ and } \mathbb{P}\left(M^{v,e} \geqslant 1 - \varepsilon\right) \leqslant \kappa_2(a_v, a_e)\varepsilon^{(a_v - a_e)/2}.$$

Proof. We abrreviate $M_{e,v} =: M$ and $M^S_{e,v} =: M^S$. Recall that $M \lhd M^S$, the last being $\beta(a,b)$-distributed with where a and b depend on $w_e(T_0)$, where T_0 is the time of the first arrival at v. Jensen's inequality for conditional expectation implies that for all convex, bounded $f : [0,1] \to \mathbb{R}$

$$\mathbb{E}[f(M)] \leqslant \mathbb{E}[f(M^S)] = \int_0^1 f(x) d\beta_{a,b}(x).$$

Now let $0 < \varepsilon < 1$. We apply this with

$$f(x) = \begin{cases} 2 - \frac{2x}{\varepsilon} & \text{if } x \leqslant \varepsilon \\ 0 & \text{else.} \end{cases} \text{ and } g(x) = \begin{cases} 0 & \text{if } x \leqslant 1-\varepsilon \\ 2 - \frac{2-2x}{\varepsilon} & \text{else} \end{cases}$$

More importantly, $\mathbb{1}_{[0,\varepsilon]} \leqslant f \leqslant 2$ and $\mathbb{1}_{[1-\varepsilon,1]} \leqslant g \leqslant 2$ and hence

$$\mathbb{P}(M \leqslant \varepsilon) \leqslant \frac{2}{\beta(a,b)} \int_0^\varepsilon x^{a-1}(1-x)^{b-1} dx \leqslant \frac{2}{a\beta(a,b)} \varepsilon^a \text{ and} \tag{3.33}$$

$$\mathbb{P}(M \geqslant 1-\varepsilon) \leqslant \frac{2}{\beta(a,b)} \int_{1-\varepsilon}^1 x^{a-1}(1-x)^{b-1} dx \leqslant \frac{2}{b\beta(a,b)} \varepsilon^b. \tag{3.34}$$

Note that $\frac{a_e}{2} \leqslant a \leqslant \frac{a_e+1}{2}$ and $\frac{a_v-a_e}{2} \leqslant b \leqslant \frac{a_v-a_e+1}{2}$. Therefore we may bound (3.33) and (3.34) by

$$\frac{2}{a\beta(a,b)} \varepsilon^a \leqslant \frac{2}{\frac{a_e}{2}\beta(\frac{a_e+1}{2}, \frac{a_v-a_e+1}{2})} \varepsilon^{\frac{a_e}{2}} =: \kappa_1(a_v, a_e)\varepsilon^{\frac{a_e}{2}} \text{ and}$$

$$\frac{2}{b\beta(a,b)} \varepsilon^a \leqslant \frac{2}{\frac{a_v-a_e}{2}\beta(\frac{a_e+1}{2}, \frac{a_v-a_e+1}{2})} \varepsilon^{\frac{a_v-a_e}{2}} =: \kappa_2(a_v, a_e)\varepsilon^{\frac{a_v-a_e}{2}}.$$

Obviously, since the β-function is continuous, κ_1 and κ_2 are continuous in both arguments. □

Lemma 3.6.10. *Let G be a finite graph and let $(a_e)_{e \in E(G)} \subset C$ for a compact set $C \subset (0,\infty)$. Then there are constants a_G, κ_G so that for all edges e, f*

$$\mu(W_e \leqslant \varepsilon W_f) \leqslant l\kappa\varepsilon^{a_G/2l}$$

where l is length of a shortest path between either ends of e and f.

Proof. Let $y = (x_0, \ldots, x_{l+1})$ be an arbitrary path in G, denote by $(e = e_0, \ldots, e_l = f)$ the corresponding edges.

$$a_G = \inf_{e \in E(G)} a_e > 0, \kappa_G = \sup_{v \in V(G), e \in \delta_G(v)} \kappa_1(a_e, a_v) < \infty.$$

The last estimate holds since $a_v \leqslant \sup_{e\in E(G)} Ka_e$ and by continuity of κ_1. For all $\varepsilon > 0$

$$\mu\left(W_e \leqslant \varepsilon W_f\right) \leqslant \mu\left(\exists i \in \{0,\dots,l\} : W_{e_{i+1}} \leqslant \varepsilon^{1/l} W_{e_i}\right) \leqslant$$
$$\sum_{i=i}^{l} \mu(W_{e_{i+1}} \leqslant \varepsilon^{1/l} W_{e_i}) \leqslant \sum_{i=1}^{l} \mu(W_{e_{i+1}} \leqslant \varepsilon^{1/l} W_{v_i}) \leqslant l\kappa_G \varepsilon^{a_G/2l}.$$

□

The tightness result from the last lemma leads to construction of a mixing measure in the transient case.

Theorem 3.6.11. *Let G be a locally finite graph and let G_n be an increasing sequence of finite subgraphs of G with limit G (for all $e \in E(G)$ there exists $N \in \mathbb{N}$ so that $e \in E(G_n)$ for all $n \geqslant N$). Define*

$$\tilde{W}_e = \frac{W_e}{W_{e*}} \tag{3.35}$$

where e^* is the first edge traversed. Denote by $\mu^{(n)}$ the unique mixing measure on G_n in terms of $\tilde{W}_e$. There is a subsequence $\mu_{k(n)}$ so that for all finite $F \subset E(G)$ the law of $(\tilde{W}_e)_{e\in E(G)}$ converges weakly to a distribution μ^* on $(0,\infty)^F$.

Proof. Fix $n \in \mathbb{N}$. Let $e \in E(G_n)$ and y an (e^*, e)-path of length l. Choose $k_0 = k_0(n)$ so that for each vertex x in y the edge set $\delta_G(x)$ belongs to G_n. Choose $\kappa := \kappa_{G_{k_0}}$ and $a := a_{G_{k_0}}$ according to lemma (3.6.10), by which we obtain

$$\mu_k(\tilde{W}_e \leqslant \varepsilon) \leqslant l\kappa\varepsilon^{a/2l} \qquad \text{and} \qquad \mu_k(\tilde{W}_e \geqslant \varepsilon^{-1}) \leqslant l\kappa\varepsilon^{a/2l}.$$

for all $k > k_0$. Hence the measures $(\mu_k)_{k\geqslant k_0}$ are tight on $(\tilde{W}_e)_{e\in E(G_n)}$ and contain a weakly convergent subsequence. Since the set of finite connected subgraphs of G is (by local finiteness of G) countable we obtain a diagonal sequence which converges weakly to a a measure μ^*. □

3.6.3 A Mixing Measure Always Exists

Now that we have constructed a measure we only need to show that it fulfills the desired properties.

Proof of Theorem 3.6.7. Let μ^* be a limit of a weakly convergent diagonal sequence from the preceding lemma. Let $y = (x_0, \dots, x_l)$ be a finite path in G. For k sufficiently big we obtain

$$
\begin{aligned}
P\left((X_0,\dots,X_l)=y\right) &= \int_{(0,\infty)^{E(G_k)}} \prod_{i=0}^{l-1} p(x_i,x_{i+1})\mu_k((0,\infty)^{E(G_k)}) \\
&= \int_{(0,\infty)^{E(G)}} \prod_{i=0}^{l-1} p(x_i,x_{i+1})\mu^*\left((0,\infty)^{E(G)}\right) \qquad (3.36)
\end{aligned}
$$

where $p(\cdot,\cdot)$ denotes the transition probabilites uniquely determined by $(\tilde{W}_e)_{e\in E(G_k)}$ or $(\tilde{W}_e)_{e\in E(G)}$. The last equality needs to hold since the integrands are continuous functions of the weights. (3.36) is the representation in terms of Markov chains. Since the the edge weights are almost surely strictly positive these Markov chains are irreducible. Reversibility follows from the representation in terms of weights. □

We finish this section with a powerful result following from the theory of Markov chains.

Corollary 3.6.12. *Let G be a locally finite, connected graph, let $v \in V(G)$. For LRRW the following statements are equivalent:*

1. *$LRRW(G,a)$ visits v infinitely often almost surely.*
2. *$LRRW(G,a)$ visits all vertices infinitely often almost surely.*

Proof.

1. "$2 \Rightarrow 1$"
   Trivial.
2. "$1 \Rightarrow 2$"
   By Theorem 3.6.1 *LRRW* is a mixture of irreducible Markov chains, denote the mixing measure by μ^*. Since v is visited infinitely often almost surely by $LRRW(G,a)$, μ^* needs to put mass 1 on the set $\mathcal{P}_r$ of recurrent Markov chains. Otherwise, if $\mu^*(\mathcal{P}_r) < 1$ then

   $$\mathbb{P}(LRRW(G,a) \text{ visits } v \text{ infinitely often}) = \mu(\mathcal{P}_r) < 1$$

   would hold. Visiting v infinitely often would not be an almost sure event, a contradiction to the first statement in this corollary.

□

Chapter 4

Results on Recurrence and Transience

4.1 The Dichotomy of Recurrence and Transience on Trees

For Markov chains it is clear what positive recurrence means. For *LRRW* let us make a note on this.

Definition 4.1.1. *We say that LRRW is almost surely transient if its mixing measure puts mass* 1 *on transient Markov chains.*
We say that LRRW is almost surely (positive) recurrent if its mixing measure puts mass 1 *on (positive) recurrent Markov chains.*

It is quite intuitive that for all locally finite graphs *LRRW* is either almost surely recurrent or almost surely transient, but to our awareness no proof has been given yet. For locally finite trees T, however, it is not difficult to prove a dichotomy between almost sure recurrence and almost sure transience. Note first that the mixing measure μ^* given as the product of jointly independent Dirichlet distributions $\prod_{v \in V(T)} D_v$ **does** represent $LRRW(T, a)$. It may or may not be unique, for the following it does not matter. Recurrence is a measurable event, it corresponds directly to a set $\mathcal{P}_r$ of recurrent Markov chains. Suppose that $0 < \mu(\mathcal{P}_r) < 1$. Then there is a vertex v so that the projection $\mathcal{P}_v$ of $\mathcal{P}_r$ to $(p(v, w))_{w \in \mathcal{N}_T(v)}$ is a set with measure $\mu(D_v \in \mathcal{P}_v) < 1$. But any Markov chain in $\mathcal{P}_r$ is recurrent for almost all values of D_v. Thus $\mathcal{P}_r$ contains the entire image of D_v a contradiction of $\mu(D_v \in \mathcal{P}_v) < 1$.

For regular trees T_{K+1} there is a phase transition for equal initial weights $a_e = a$. The approximate critical values for different K are given in Table 6.1. If not mentioned differently **each result in this section is due to Pemantle**,

[17]. In his notation the reinforcement function is assumed to be

$$f(n) = 1 + \Delta n$$

for $\Delta \in \mathbb{R}^+$. In fact the theorem is not limited to fixed regular trees but Galton-Watson trees, assuming i.i.d. distribution of the number of children of each vertex. We will only formulate it for fixed $(K+1)$-regular trees.

Theorem 4.1.2. *For LRRW on an infinite $(K+1)$-regular tree $T_{K+1} = (V, E)$, there exists a constant $a_0(K)$ so that the process is almost surely positive recurrent if $a < a_0$ and transient if $a > a_0$.*

We denote by x_0 the initial vertex, w.l.o.g. x_0 is the root of the tree. For all $v \in V(T) - x_0$ we will denote by $Par(v)$ and $Par^2(v)$ its parent and grandparent and by $Ch(v)$ the set of its children. For $v, w \in V(T)$ we write $v < (\leqslant)\ w$ if w is a descendant of v (a descendant of v or $w = v$). The density for the transition probabilities from each vertex v except x_0 is given by

$$f_v(p_0^v, p_1^v, \ldots, p_K^v) = \frac{\Gamma\left(\frac{(m+1)a+1}{2}\right)}{\Gamma\left(\frac{a+1}{2}\right)\Gamma\left(\frac{a}{2}\right)^K}\left(1 - \sum_{i=1}^K p_i\right)^{\frac{a+1}{2}} (p_1^v \cdots p_K^v)^{\frac{a}{2}}$$

$$p_i^v \in [0,1], i = 0, \ldots, K, \quad \sum_{i=0}^K p_i^v = 1$$

where $K := \deg_T(v) - 1$, is the number of children of v for $v \neq x_0$ and $p_0 = p(v, Par(v))$. Recurrence and transience do not depend on a finite number of vertices, so we may as well assume that the transition probabilities for all vertices are distributed like above. We denote by $D(Par(v), v)$ and $D(Par(v), Par^2(v))$ the projections of $D_{Par(v)}$ to p_0 and p_i for some $i \in \{1, \ldots, K\}$, respectively. Define

$$\phi(v) := \frac{D(Par(v), v)}{1 - D(Par(v), Par^2(v))}.$$

and

$$\psi(r) = \inf\{e^{-rt}\mathbb{E}[\phi^t] : t \in \mathbb{R}\}.$$

The $\phi(v)$ are not jointly independent, being obvious when looking at siblings v_1, v_2. However, if a set $V' \subset V(T)$ does not contain more than one child per vertex, joint independence of $\phi(v)_{v \in V'}$ holds true since the D_v are independent.

Theorem 4.1.3. *Abbreviate* $\tilde{\psi}(r) := \psi(\ln(r))$.

1. *If* $\mathbb{E}[\ln(\phi)] \geqslant 0$ *then* $LRRW(T, a)$ *is almost surely transient.*
2. *If* $\mathbb{E}[\ln(\phi)] < 0$ *and* $\sup\{Kr\tilde{\psi}(r) : 0 < r \leqslant 1)\} < 1$ *then* $LRRW(T, a)$ *is almost surely positive recurrent.*
3. *If* $\mathbb{E}[\ln(\phi)] < 0$ *and* $\sup\{Kr\tilde{\psi}(r) : 0 < r \leqslant 1)\} > 1$ *then* $LRRW(T, a)$ *is almost surely transient.*
4. *If* $1 \leqslant \mathbb{E}[\phi] \leqslant \infty$ *then the* sup *in 2 and 3 occurs at* $r = 1$.

4.1.1 The Transient Case

Parts 1 and 3 are reduced to the following lemma, constructing a transient Markov chain.

Lemma 4.1.4. *By abuse of notation we define for a Markov chain* X *with transition matrix* $P = (p(v, w))_{v,w \in V(T)}$

$$\phi(v) := \phi(v|P) = \frac{p(Par(v), v)}{1 - p(Par(v), Par^2(v))}. \tag{4.1}$$

Let $k \in \mathbb{N}^*, L \in \mathbb{R}^+, \delta > 0, r \in (0, 1]$ *be some fixed constants. Denote by* T_i *the set of vertices at distance* i *from* x_0. *Suppose we can find a nonempty set* $S \subset V(T)$ *with the following properties, writing* $S_i := S \cap T_{ik}$.

1. $v \in S_{i+1} \Rightarrow \exists v' \in S_i : v' \leqslant v$
2. $v \in S_i \Rightarrow |\{w \in S_{i+1} : w < v\}| \geqslant r^{-k}$
3. *For all paths* $v = v_0 < v_1 < \cdots < v_k$ *with* $v_0 \in S_i$ *and* $v_k \in S_{i+1}$
$$\sum_{i=1}^{k} \ln(\phi(v_i)) \geqslant k \ln(r) + \delta \tag{4.2}$$
4. $\phi(v)^{-1} \leqslant L$ *for all* $v \in S$

Then X *is transient.*

To understand the intuition behind this look at the case $r = 1$. Then a set S contains a path $v_0, v_1, \ldots$ with $v_i < v_{i+1}$. The sequence $f(v_i) = \prod_{v < v_i} \phi(v)^{-1}$ is summable by the properties 3 and 4 of S. Restricting the process to the path the expected number of visits to v_0 is finite and thus the process needs to be transient.

Proof. Define a function $s : V(T) \to [0,1]$ by

$$s(v) = \begin{cases} \dfrac{|\{w \in S_{i+1} : v \leqslant w\}|}{|\{w \in S_{i+1} : Par(v) \leqslant w\}|} & \text{if } v \in S \\ 0 & \text{else} \end{cases} \tag{4.3}$$

where $v \in T_j, ik < j \leqslant (i+1)k$. Note that for each $v \in S, \sum_{w \in Ch(v)} s(w) = 1$. We may view the pair (S, s) as a multistage experiment. Define recursively

$$\begin{aligned} t &: V(T) \to \mathbb{R}^+, t(x_0) = 1 \\ t(v) &= \frac{s(v)}{\phi(v)} t(Par(v)). \end{aligned}$$

and

$$M(v) := \sum_{w \leqslant v} t(w).$$

Now let $\tilde{X}, \tilde{X}_0 = x_1 \neq x_0$ be the Markov chain with transition matrix

$$\tilde{p}(v, w) = \begin{cases} p(v, w) & v \neq x_0 \\ 0 & v = x_0, w \neq x_0 \\ 1 & v, w = x_0, \end{cases} \tag{4.4}$$

i.e. as X but stopped at the first arrival at x_0. We show that $M(\tilde{X}_i)$ is a bounded martingale.

$$\begin{aligned} &\mathbb{E}[M(\tilde{X}_{i+1})|\tilde{X}_i] = \\ =&D(\tilde{X}_i, Par(\tilde{X}_i))M\left(Par(\tilde{X}_i)\right) + \sum_{w \in Ch(\tilde{X}_i)} D(\tilde{X}_i, w)M(w) = \\ =&M(\tilde{X}_i) + D(\tilde{X}_i, Par(\tilde{X}_i))\left(-t(\tilde{X}_i) + \sum_{w \in Ch(\tilde{X}_i)} \phi(w)t(w)\right) = \\ =&M(\tilde{X}_i) + \frac{p(\tilde{X}_i, Par(\tilde{X}_i))}{t(\tilde{X}_i)}\left(-1 + \sum_{w \in Ch(\tilde{X}_i)} s(w)\right) = M(\tilde{X}_i) \end{aligned}$$

To show that M is bounded, consider the case $v_k \in S_i$. Let $v_0 \in S_{i-1}$, $v_0 < v_k$. By properties 2 and 4 in Lemma 4.1.4

$$\prod_{i=1}^{k} s(v_i) \leqslant r^k \text{ and } \prod_{i=1}^{k} \phi(v_i)^{-1} \leqslant e^{-\delta} r^{-k}.$$

Thus $t(v_k) \leqslant t(v_0)e^{-\delta}$ and by induction on i, $t(v_k) \leqslant e^{-i\delta}$. Since $t(v)$ decreases exponentially with the level of v, $M(v)$ is uniformly bounded in T. Hence by the martingale convergence theorem $M(\tilde{X}_i)$ converges almost surely to a random limit M_∞ with $\mathbb{E}[M_\infty] = \mathbb{E}[M(x_1)] > 1$. Thus $\mathbb{P}(M_\infty = 1) < 1$ needs to hold true. More importantly, the probability of never visiting x_0 is positive. Thus the Markov chain is transient. □

We do not need to show that 1 and 3 are sufficient conditions for a set S like in the last lemma, we only need to show that the existence of S has **positive** probability. By deleting subtrees that do not contain elements in S it becomes obvious that branching processes are a highly useful concept to eventually get a set like S. For the following two theorems we will not give a proof. The first one is due to Chernoff, [2]. It estimates probabilities of large deviations. We apply the theorem with $X_i = \ln(\phi)$ to construct a set S as in the last lemma. The second one is a general result for Galton-Watson trees and is due to Harris, [7].

Theorem 4.1.5 (Chernoff, [2])**.** *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of i.i.d. random variables. Let $S_n := X_1 + \cdots + X_n$. For $r \in \mathbb{R}$ let again*

$$\psi(r) := \inf\{e^{-rt}\mathbb{E}[e^{tX_1}] : t \in \mathbb{R}\}.$$

Let $r > \mathbb{E}[X_n] \geqslant -\infty$. Then

$$\mathbb{P}(S_n \geqslant nr) \leqslant \psi(r)^n \tag{4.5}$$

and for any $\psi_0 < \psi(r)$

$$\lim_{n\to\infty} \psi_0^{-n}\mathbb{P}(S_n \geqslant nr) = \infty. \tag{4.6}$$

Further, ψ is continuous in r and strictly decreasing between $\mathbb{E}[X_1]$ and $\operatorname{essup} X_1$.

Theorem 4.1.6 (Harris, [7])**.** *Let B be a branching process. Each vertex bears a random number I of children assumed to be i.i.d. and equal to i with probability q_i. Suppose that $1 < \lambda := E[I] < \infty$, so the probability of non-extinction is some positive value $0 < b < 1$. Let $\epsilon > 0, 0 < \lambda_0 < \lambda$. Then*

$$\lim_{n\to\infty} \mathbb{P}(\textit{size of the n-th generation} < \epsilon\lambda_0^n) = b.$$

We call a tree d-infinite if it contains a subtree in which each vertex has at least d children. We say that a vertex v has a (d, n)-subtree if $n = 0$ or v has at least d children, each having a $(d, n-1)$-subtree. Suppose that B is a branching process with generating function $f(x)$ and C the branching process with generating function $\bar{f}(x) = f(r + (1-r)x)$. Comparing f and $\bar{f}$ it becomes clear that B and C are structurally not much different. The only difference is that in C already existing vertices in a new generation are deleted with probability r.

Lemma 4.1.7. *Suppose that for C, the probability of a vertex having at least d children, is at least $1-r$. Then B is d-infinite with probability at least $1-r$.*

Proof. We show by induction that

$$\mathbb{P}(C \text{ has a } (d,n) - \text{subtree}) \geqslant 1-r.$$

For $n=0$ this is trivial. The probability of a vertex having at $(d,n+1)$-subtree is the probability of having at least d children assuming that those who do not have a (d,n)-subtree are deleted. By the induction hypothesis, children will be deleted with probability at most r. □

Lemma 4.1.8. *For a branching process B denote by $B^{(k)}$ the process for which*

$$B_n^{(k)} := B_{kn}.$$

Let λ, λ_0 as in Lemma 4.1.6. Then there is some $k \in \mathbb{N}$ so that

$$\mathbb{P}(B^{(k)} \textit{ is } \lfloor \lambda_0^k \rfloor\textit{-infinite}) \geqslant \frac{1-b}{2}.$$

Proof. By Theorem 4.1.6 for some N sufficiently large

$$\mathbb{P}\left(\text{size of } n\text{-th generation } > \frac{4\lambda_0^n}{1-b}\right) > \frac{3(1-b)}{4}$$

holds for all $n \geqslant N$. Now given a population of size $\geqslant \frac{4\lambda^N}{1-b}$, if we kill each individual independently with probability $\frac{1+b}{2}$ for N sufficiently large

$$\mathbb{P}\left(\text{at least } \lambda_0^n \text{ of them survive}\right) \geqslant \frac{3(1-b)}{4}.$$

The last inequality is just a consequence of the law of large numbers. Now we apply Lemma 4.1.7 with $B^{(N)}$ and $r = \frac{1+b}{2}$. For the process C the probability of a vertex having at least λ_0^n children is

$$\begin{aligned}
&\mathbb{P}\left(v \text{ has at least } \lambda_0^N \text{ children}\right) \geqslant \\
\geqslant &1 - \mathbb{P}\left(\text{fewer than } \frac{4\lambda_0^N}{1-b} \text{ children are born}\right) \\
&- \mathbb{P}(\text{from more than } \frac{4\lambda_0^N}{1-b} \text{ children less than } \lambda_0^N \text{ survive}) \geqslant \\
\geqslant &1 - \frac{1-b}{4} - \frac{1-b}{4} = \frac{1-b}{2}
\end{aligned}$$

Hence $B^{(N)}$ is λ_0^N-infinite with probability of least $\frac{1-b}{2}$. □

Proof of Parts 1 and 3 of Theorem 4.1.3. Part 1 follows easily by the strong law of large numbers for $\ln(\phi)$ and Lemma 4.1.8 choosing some r sufficiently close to 1. Then for k and L sufficiently large we may use Lemma 4.1.8 to construct S. We will thus only concentrate on the more difficult case of Part 3. Fix r so that $Kr\tilde{\psi}(r) > 1$. Let $\delta_1 > \delta_2 > \delta_3 > 0$ so that

$$\tilde{\psi}(r) = \frac{1+\delta_1}{Kr} > \frac{1+\delta_2}{Kr} > \frac{1}{Kr}.$$

For shortness, write

$$T_n(\epsilon) = \{v \in T_n : \sum_{w<v} \ln(\phi(v)) > n\ln(r) + \epsilon\}.$$

By (4.5) in 4.1.5, since $\frac{1+\delta_2}{Kr} < \tilde{\psi}(r)$, for N sufficiently large and δ_0 sufficiently small

$$\mathbb{E}\left(|T_N(\delta_0))|\right) > K\left(\frac{1+\delta_2}{r}\right)^N.$$

Picking L sufficiently large we may amend this to

$$\mathbb{E}\left[\left|\{v \in T_N(\delta_0) : \phi(w)^{-1} < L \text{ for all } w \leqslant v\}\right|\right] > K\left(\frac{1+\delta_3}{r}\right)^N.$$

We now define a branching process $B^{(N)}$ with initial ancestor x_0 and elements of $(T_{kN})_{k\in\mathbb{N}}$, where $v \in T_{kN}$ has $w \in T_{(k+1)N}, w > v$ as a child if and only if

$$\sum_{v\leqslant v'\leqslant w} \ln(\phi(v')) \geqslant N\ln(r) + \delta_0 \text{ and } (\phi(v'))^{-1} < L$$

holds for all $v', v \leqslant v' < w$. By Lemma 4.1.8 the probability of the existence of a set S like in Lemma 4.1.4 is strictly positive. Since μ^* does not mix transience and recurrence the process is almost surely transient. □

4.1.2 The Recurrent Case

For the proof of Part 2 of Theorem 4.1.3 we construct an almost surely stationary distribution α_v by

$$f(v) := \prod_{x_0<w\leqslant v} \phi(w)$$
$$\alpha_{x_0} := 1, \alpha_v := (D(v, Par(v)))^{-1} f(v).$$

Note that for $w \in Ch(x_0)$ the expression $\phi(w)$ is not defined. But we may as well add a vertex to x_0 and define it as the root instead. This does not affect the process of being recurrent or transient. With this definition, α satisfies for all v and $w \in Ch(v)$

$$\alpha_v D(w, v) = \alpha_w D(v, w).$$

To show that $\sum_{v \in V(T)} \alpha_v < \infty$ we make use of the following lemma:

Lemma 4.1.9. *Let $k > 1$, $0 < r < 1$. Suppose that $\mathbb{E}[\ln(\phi)] < 0$. Define*

$$A_n(r) := \{v \in T_n : \alpha_v \geqslant r^n\}.$$

Then

$$\mathbb{P}(|A_n(r)| \geqslant (Kk\tilde{\psi}(r))^n \textit{ for infinitely many } n) = 0.$$

Proof. Define

$$A'_n(r) = \{v \in T_n : f(v) \geqslant r^n\}.$$

For each $v \in T_n$, by statement (4.5) of Theorem 4.1.5

$$\begin{aligned} &\mathbb{P}\left(f(v) \geqslant r^n\right) = \\ =&\mathbb{P}\left(\sum_{x_0 < w \leqslant v} \ln(\phi(w)) \geqslant n \ln(r)\right) \leqslant (\tilde{\psi}(r))^n. \end{aligned}$$

Using linearity of expectation and noting that $|T_n| = K^n$ and $k > 1$

$$\mathbb{E}[|A'_n(r)|] \leqslant (K\tilde{\psi}(r))^n \tag{4.7}$$

and thus

$$\mathbb{E}\left[\sum_{n \in \mathbb{N}} \frac{|A'_n(r)|}{(Kk\tilde{\psi}(r))^n}\right] < \infty. \tag{4.8}$$

We conclude that at most finitely many summands in (4.8) are bigger than 1 and so

$$\mathbb{P}(|A'_n(r)| \geqslant (Kk\tilde{\psi}(r))^n \text{ for infinitely many } n) = 0. \tag{4.9}$$

We claim that (4.9) still holds true replacing A'_n by A_n. Suppose that

$$A(r) := [|A_n(r)| \geqslant (Kk\tilde{\psi}(r))^n \text{ f.i.m. } n]$$

occurs with positive probability. Since $\tilde{\psi}$ is continuous and decreasing in $(0, \infty)$ we may choose r_1, k_1 $r > r_1 > 0$ and $k > k_1 > 1$ so that

$$\tilde{\psi}(r_1) < 1 \text{ and } k_1\tilde{\psi}(r_1) = k\tilde{\psi}(r).$$

Thus

$$A(r) = [|v \in T_n : \alpha_v \geqslant r^n| \geqslant (Kk_1\tilde{\psi}(r_1))^n \text{ for infinitely many } n].$$

Now choose k_2, $k_1 > k_2 > 1$ and choose N sufficiently large so that

$$(Kk_1\tilde{\psi}(r_1))^n \geqslant (Kk_2\tilde{\psi}(r_1))^{n+1}$$

holds for $n \geqslant N$. Since $\alpha_v = (D(v, Par(v)))^{-1}f(v)$, for all $w \in Ch(v)$, $f(w) = \alpha_v D(v,w)$ holds true. Thus the event

$$[|v \in T_n : f(w) \geqslant r^n D(v,w) \text{ for some } w \in Ch(v)| \geqslant (Kk_2\tilde{\psi}(r_1))^{n+1} \text{ f.i.m. } n]$$

has positive probability as well. Now $D(v,w)$ and $f(w)$ are not independent, but even better. For $p \in [0,1]$

$$\begin{aligned}&\mathbb{P}(D(v,w) \geqslant p \text{ for at least one } w \in Ch(v)|\alpha_v \geqslant r^n) \geqslant\\ &\mathbb{P}(D(v,w) \geqslant p \text{ for at least one } w \in Ch(v)).\end{aligned}$$

Thus the event

$$[|\{w \in T_{n+1} : f(w) \geqslant r_1^{n+1}\}| \geqslant (Kk_2\tilde{\psi}(r_1))^{n+1} \text{ f.i.m. } n]$$

has positive probability, a contradiction to (4.9). □

Proof of Part 2 of Theorem 4.1.3. We would like to show almost sure finiteness of

$$\sum_{v \in V(T)} \alpha_v = \sum_{n \in \mathbb{N}} \sum_{v \in T_n} \alpha_v.$$

Then, after normalizing, α is a stationary distribution.
Let $\sup\{Kr\tilde{\psi}(r) : r > 0\} = 1 - \delta_1$. Choose δ_2, δ_3 with $0 < \delta_3 < \delta_2 < \delta_1$. Let $l_1 \in (0,1)$ so that $\mathbb{E}[\ln(\phi)] < \ln(l_1) < \ln\left(\frac{1}{K}\right)$. Let $g : [0,1] \to \mathbb{R}$ be any function satisfying

$$g(t) < t, \quad Kt\tilde{\psi}(g(t)) < 1 - \delta_3.$$

The purpose of g is to generate a cover $\{(l_i, u_i)\}_{i \in I} \cup \{g(1), 1\}$ of $(l_1, 1]$. The family $\mathcal{O} := \{(g(t), t)\}_{t \in (0,1)} \cup \{(g(1), 1)\}$ contains a cover of $[l_1, 1]$. Since $[l_1, 1]$ is compact $\mathcal{O}$ contains a finite cover

$$\{(l_i, u_i)\}_{i=1,\ldots,k}$$

of $(l_1, 1]$ with $l_i < l_{i+1}$. Take $l_0 := 0, u_0 := l_1, l_{k+1} :=$. For all intervals $(l_i, u_i), i = 1, \ldots, k$ and for $(g(1), 1]$ we may apply Lemma 4.1.9 with $r = l_i$ and $k = \frac{1-\delta_3}{1-\delta_2}$. Thus with probability 1 there is some N so that

$$|A_n(r)| < \left(K\frac{1-\delta_3}{1-\delta_2}\tilde{\psi}(l_i)\right)^n$$

holds for all i, for all $n \geqslant N$. Hence we may bound for all $n \geqslant N$

$$\sum_{v \in T_n} \alpha_v = \sum_{i=0}^{k+1} \sum_{v \in T_n : l_i \leqslant \alpha_v \leqslant u_i} \alpha_v \leqslant$$
$$\leqslant K^n l_1^n + \sum_{i=1}^{k} u_i^n \left(K \frac{1-\delta_2}{\delta_3} \tilde{\psi}(l_1) \right)^n \leqslant (k+1)(1-\delta)^n$$

where the last equation follows by $u_i K \tilde{\psi}(l_i) < 1 - \delta_3$. Thus α is almost surely a stationary distribution. □

4.1.3 The Calculation of the Phase Transition

The calculation of $\tilde{\psi}$ might be cumbersome. However, Part 4 of Theorem 4.1.3 will imply that it is not necessary to calculate $\tilde{\psi}$ but only $\tilde{\psi}(1) = \psi(0)$ to obtain the parameter at which the phase transition occurs.

Proof of Part 4 of Theorem 4.1.3 and Theorem 4.1.2. We now investigate which case in Theorem 4.1.3 needs to be applied. We first establish Part 4. Suppose that $1 \leqslant \mathbb{E}[\phi] \leqslant \infty$. For $r \in (0,1]$

$$K r \tilde{\psi}(r) = \inf\{K r^{1-t} \mathbb{E}[\phi^t], t \in \mathbb{R}\}. \tag{4.10}$$

Both terms r^{1-t} and $\mathbb{E}[\phi^t]$ are increasing in t for $t > 1$ since by Jensen's inequality $\mathbb{E}[\phi^t] \geqslant \mathbb{E}[\phi]^t \geqslant \mathbb{E}[\phi]$. Thus the infimum on the right hand side occurs for $t \leqslant 1$. But for $t \leqslant 1$ the expression in (4.10) is increasing in r so the supremum of the left hand side must occur at $r = 1$. Now we first determine $a_0(K)$ supposing that $E[\phi] \geqslant 1$. We will show later that this holds true anyway in a neighbourhood of $a_0(K)$. Since $(p_0, p_1, p_2 + \cdots + p_K)$ is Dirichlet-distributed with parameters $\frac{a+1}{2}$, $\frac{a}{2}$ and $\frac{(K-1)a}{2}$,

$$\psi(0) = \inf\left\{ \mathbb{E}\left[\left(\frac{p_1}{p_0} \right)^t \right] \right\} = \tag{4.11}$$
$$= \frac{\Gamma\left(\frac{(K+1)a}{2}\right)}{\Gamma\left(\frac{a+1}{2}\right)\Gamma\left(\frac{a}{2}\right)\Gamma\left(\frac{(K-1)a}{2}\right)} \cdot$$
$$\inf_t \left\{ \int\int p_0^{\frac{a+1}{2}-t-1} p_1^{\frac{a}{2}+t-1} (1-p_0-p_1)^{\frac{(K-1)a}{2}-1} dp_0 dp_1 \right\} =$$
$$= \inf_t \frac{\Gamma\left(\frac{a+1}{2}-t\right)\Gamma\left(\frac{a}{2}+t\right)}{\Gamma\left(\frac{a+1}{2}\right)\Gamma\left(\frac{a}{2}\right)}. \tag{4.12}$$

This expression is convex in t and symmetric about $\frac{1}{4}$ which is thus the minimizer. We define $a_0(K)$ for $K = 2, 3, \dots$ as the unique positive solution of

$$f_\Gamma(a) := \frac{\left(\Gamma\left(\frac{a}{2} + \frac{1}{4}\right)\right)^2}{\Gamma\left(\frac{a+1}{2}\right)\Gamma\left(\frac{a}{2}\right)} = \frac{1}{K}. \tag{4.13}$$

To see that a solution exists for $K \geqslant 2$ note that f_Γ is continuous in $(0, \infty)$ and

$$\lim_{a\to 0^+} f_\Gamma(a) = 0, \quad \lim_{a\to\infty} f_\Gamma(a) = 1.$$

To see that it is unique observe that the integrand for $t = \frac{1}{4}$ is strictly decreasing in a. For $a_0 := a_0(K)$ we have

$$\psi(0) = \begin{cases} \frac{\Gamma\left(\frac{a_0-1}{2}\right)\Gamma\left(\frac{a_0}{2}+1\right)}{\Gamma\left(\frac{a_0+1}{2}\right)\Gamma\left(\frac{a_0}{2}\right)} = \frac{a_0}{a_0+1} & \text{if } a_0 > 1 \\ \infty & \text{else} \end{cases}.$$

It suffices to show that in a neighbourhood of $a_0(K)$, $\mathbb{E}[\phi] \geqslant 1$. f_Γ is strictly decreasing. Taking $a_0(K) = 0.9$ we get

$$f_\Gamma(0.9) > 0.83 > \frac{1}{2} \geqslant \frac{1}{K},$$

so $a_0(K) < 1$ for $K \geqslant 2$ and hence in a neighborhood of $a_0(K)$, $\mathbb{E}[\phi] = \infty$. Taking $K = 1$ results in the special case of $LRRW(\mathbb{N}, a)$. For $K = 1$ equation (4.13) has no solution. Thus $LRRW(N, a)$ is recurrent for all initial weights a. Table 6.1 shows numerical approximations of the critical values $a_0(K)$ for $K = 1, 2, \dots, 15$. $\square$

4.2 Almost Sure Positive Recurrence for Bounded Degrees and Small Initial Weights

Until March 2012 little was known about return probabilities of *LRRW* in general graphs. Angel, Crawford and Kozma, [1], managed to prove the following.

Theorem 4.2.1. *Let $K \in \mathbb{N}$. Then there exists a constant $a_0(K) > 0$ so that on any graph G with vertex degrees bounded by K, $LRRW(G, a)$ is almost surely positive recurrent for all $a \leqslant a_0$ componentwise.*

The result is somewhat intuitive. For instance, it is a well-known fact that simple random walk is transient on $\mathbb{Z}^3$. Hence, if the initial weights a are sufficiently big it is natural to conjecture that the process behaves similar to simple random walk and is thus transient. However, for small initial weights, an increment of 1 for an edge has a much bigger impact and makes it much more likely to return to the initial (or any other) vertex. This conjecture has been proved rigorously by Sabot and Tarrès, [20]. For $d \geqslant 3$, $LRRW(\mathbb{Z}^d, a_e = a)$ is recurrent for sufficiently small and transient for sufficiently large initial weights. For $d \in \{1, 2\}$ recurrence always holds, the case of $d = 1$ being a consequence of Theorem 4.1.3. The proof by Sabot and Tarrés relies on a model called the Vertex Reinforced Jump Process. The approach by Angel, Crawford and Kozma is completely different.
We give a short overview of the proof strategy. Like in Theorem 4.1.3 we construct a stationary distribution. To do so we want to show that for almost all edges e the weight W_e fulfills $W_e \leqslant c^{d(e,x_0)}$ where $d(e, x_0)$ is the distance of e from the initial vertex and $c \in (0, 1)$. This way assuming bounded vertex degrees is crucial, we need to bound the number of edges at a certain distance l to the initial vertex which is of order at most K^l, explaining the role of K. The implication that this bound holds for almost all weights will be a consequence of Markov's inequality. Like in the proof of Theorem 3.6.11 we will approximate the weights on a sequence of finite subgraphs. The notion of convex order will not play a role, instead the path by which the vertices are visited first is considered. We replace the dependent normalized weights by a set of independent random variables, being the most unconventional part of the proof. If not mentioned differently, the results in this chapter are due to Angel, Crawford and Kozma, [1].

4.2.1 Bounding the Edge Weights

For the next results we will make use of the mixing measure μ on the space $(0, \infty)^{E(G)}$. Normalizing $\bar{W}_e := \frac{W_e}{W_{x_0}}$ for the initial vertex of the process x_0 does not have any impact on the chosen Markov chain. For two vertices u, v denote by $d(u, v)$ the length of a shortest path from u to v. Analogously, define $d(v, e) := \min_{x \in e} d(v, x)$. We state exponential decay in the following way.

Theorem 4.2.2. *Let G be an infinite connected graph satisfying $\sup_{v\in V(G)} \deg_G(v) \leqslant K$ for some $K \in \mathbb{N}$. Let $s \in (0, \frac{1}{4})$. Let x_0 be the initial vertex and e_1 the first edge traversed. Then there exist constants $a_0 := a_0(s, K)$ and C so that*

$$\mathbb{E}[\bar{W}_e^s] \leqslant \mathbb{E}\left[\left(\frac{\bar{W}_e}{\bar{W}_{e_1}}\right)^s\right] \leqslant 2K\left(C\sqrt{a_0}\right)^{d(e,x_0)}. \tag{4.14}$$

It would be nice to prove Theorem 4.2.2 for $s = 1$ since fractional s may only complicate any computation. However, even if G is a tree the ratio of edge weights does not need to have finite first moment. The proof of Theorem 4.2.2 will make use of directed edges instead of undirected ones, substituting each undirected edge in G by two directed ones. For a directed edge $e^+ = (u, v)$ we will denote by $e^- = (v, u)$ its reversal. Now let $e^+ = (u, v)$ be traversed by the walk. For simplicity of notation we suppose that the path in the following construction has length l. Let $e_l^+ = (u_l, u)$ be the edge through which u is visited first and $e_{l-1}^+ = (u_{l-1}, u_l)$ the edge through which u_l is visited first and so on. This results in a path $Y_\Gamma^+ := (x_0 = u_1, u_2, \ldots, u) = (e_1^+, \ldots, e_l^+)$ starting in the initial vertex and ending in u. On the fact that, in the recurrent case, this path determines the distribution of $M_{e,v}^S$ and $M^{e,v}$. Clearly, by construction, each edge $e_1^+, \ldots, e_l^+$ is traversed by *LRRW* before its corresponding inverse e_i^-. We will call $\gamma^+ = (x_0 = u_1, \ldots, u_l)$, a random variable itself, the path of domination. For a directed edge e^+ we will denote by D_γ the event that γ is the path of domination of e^+.

Since $\frac{\tilde{W}_e}{\tilde{W}_{e_1}} = \prod_{i=1}^{l-1} \frac{W_{e_{i+1}}}{W_{e_i}}$, bounding each of the factors will be sufficient, so we would like to estimate $\frac{\tilde{W}_{e_i}}{\tilde{W}_{e_{i+1}}}$. For shortness we will write $e := e_{i+1}^+, f = e_i^-$ throughout the remaining part of this chapter. We introduce two random variables $N(e)$ and $N(f)$ as follows. If e is traversed before f then take $N(e)$ to be the number of times that e is traversed before the first appearance of f and $N(f) = 1$, else vice versa. We define

$$\tilde{R}(e) := \frac{N(e)}{N(f)}, \quad R(e) = \frac{W_e}{W_f}. \tag{4.15}$$

Hence we are estimating the ratio of weights R by the ratio of numbers of directed edge traversals $\tilde{R}$.

Now let g be an arbitrary undirected edge in $E(G)$. Denote by Y_Γ the set of paths of domination, by Y_Γ^g the set of paths of domination terminating in g. The identity

$$\mathbb{E}\left[\left(\frac{W_g}{W_{e_1}}\right)^s\right] = \sum_{\gamma\in Y_\Gamma^g} \mathbb{E}\left[\left(\frac{W_g}{W_{e_1}}\right)^s \mathbb{1}_{D_\gamma}\right]$$

holds under recurrence and in particular if G is finite. Note that without the assumption of recurrence we cannot ensure that the random variable $\frac{W_g}{W_{e_1}}$ is well-defined. For any fixed event D_γ we telescope

$$\frac{W_g}{W_{e_1}} = \prod_{e' \in \gamma - e_1} R(e') = \prod_{e' \in \gamma - e_1} \frac{R(e')}{\tilde{R}(e')} \prod_{e' \in \gamma - e_1} \tilde{R}(e').$$

Applying the Cauchy-Schwarz inequality yields

$$\begin{aligned} &\mathbb{E}\left[\left(\frac{W_g}{W_{e_1}}\right)^s \mathbb{1}_{D_\gamma}\right] \leqslant \\ \leqslant &\mathbb{E}\left[\prod_{e' \in \gamma - e_1} \left(\frac{R(e')}{\tilde{R}(e')}\right)^{2s} \mathbb{1}_{D_\gamma}\right]^{\frac{1}{2}} \mathbb{E}\left[\prod_{e' \in \gamma - e_1} \tilde{R}(e')^{2s} \mathbb{1}_{D_\gamma}\right]^{\frac{1}{2}}. \end{aligned}$$

It suffices to establish sufficiently small bounds for each of the terms on the right hand side of the last equation. We will show this in the two following independent lemmas. The main idea in both is to ignore the event D_γ at some suitable step.

Lemma 4.2.3. *For any instance (G, a) so that $LRRW(G, a)$ is recurrent, any inital vertex x_0, any $\gamma \in Y_\Gamma$ with lenght l and any $s \in (0, 1)$ there exists a constant $c(s)$ depending only on s so that*

$$\mathbb{E}\left[\prod_{e' \in \gamma - e_1} \left(\frac{R(e')}{\tilde{R}(e')}\right)^{s} \mathbb{1}_{D_\gamma}\right] \leqslant c(s)^{l-1}. \tag{4.16}$$

Lemma 4.2.4. *Let (G, a) be a recurrent instance, $\deg_G(v) \leqslant K$, $a \leqslant a_0$ for a constant a_0. For any initial vertex x_0, any $\gamma \in Y_\Gamma$ with length l nd any $s \in (0, \frac{1}{2})$, there exists a constant $C(s, K)$ depending only on s and K so that*

$$\mathbb{E}\left[\prod_{e' \in \gamma - e_1} \tilde{R}(e')^{s} \mathbb{1}_{D_\gamma}\right] \leqslant (C(s, K) a_0)^{l-1}. \tag{4.17}$$

Proof of Lemma 4.2.3. The first thing to note is, since we want to show exponential decay, we do not really care about exponentially growing terms like for instance K^l, we will still be able to keep the expectation bounded in (4.17) by adjusting a_0. Thus the first step is to ignore the event D_γ in (4.16). It is straightforward that

$$\mathbb{E}\left[\prod_{e' \in \gamma - e_1} \left(\frac{R(e')}{\tilde{R}(e')}\right)^{s} \mathbb{1}_{D_\gamma}\right] \leqslant \mathbb{E}\left[\prod_{e' \in \gamma - e_1} \left(\frac{R(e')}{\tilde{R}(e')}\right)^{s}\right].$$

Let again $e := e_{i+1}^+, f := e_i^-$. Observe that recurrence is still assumed and that the definition of $\tilde{R}(e)$ is not tied to D_γ, so $R(e)$ and, especially, $\tilde{R}(e)$ are well-defined. Now let $\mathbb{W}$ be one realization of the edge weights. Conditioning on $\mathbb{W}$ makes $R(e)$ degenerate and $\tilde{R}(e)_{e\in\gamma}$ independent, the second being a consequence of the Markov property. We will show that

$$\mathbb{E}\left[\left(\frac{R(e)}{\tilde{R}(e)}\right)^s \middle| \mathbb{W}\right] \leqslant C(s)$$

holds uniformly in $\mathbb{W}$. Note that the expression is almost surely (w.r.t. $\mathbb{W}$) well-defined, for instance by Lemma 3.6.10. Denote by v the vertex incident to e and f. We are only interested in transversals of e and f from v. Neither transversals of other edges from v nor transversals of e or f in the opposite direction have any impact on $R(e)$ and $\tilde{R}(e)$, given $\mathbb{W}$. Since v is visited infinitely often anyway we may as well assume $\deg_G(v) = 2$. Denote the probabilities of transversals of e and f from v w.r.t. $\mathbb{W}$ by p and q, respectively. Of course, $p + q = 1$. But $\frac{1}{\tilde{R}(e)} = \frac{N(f)}{N(e)}$ and so

$$\mathbb{E}\left[\left(\frac{R(e)}{\tilde{R}(e)}\right)^s \middle| \mathbb{W}\right] = \left(\frac{p}{q}\right)^s \left(\sum_{n\geqslant 1} n^s q^n p + \sum_{n\geqslant 1} n^{-s} p^n q\right).$$

This expression converges for all $p, q \in (0, 1)$. The first sum is the (fractional) s-moment of a geometric random variable with parameter p and is of order $p^{-s}q$. The second term is the $-s$-moment of a geometric random variable with parameter q and is of order pq^s. Those two facts can be proved using the representation of the polylogarithm function

$$\mathrm{Li}_s(z) = \sum_{n\geqslant 1} \frac{z^k}{k^s} = \frac{1}{\Gamma(s)} \int_0^\infty \frac{t^{s-1}}{\frac{e^t}{z} - 1} dt, \mathrm{Re}(s) > 0$$

and

$$\mathrm{Li}_{s-1}(z) = z \frac{\partial \,\mathrm{Li}_s(z)}{\partial z},$$

see for instance Cvijović, [4]. Together with the factor $\left(\frac{p}{q}\right)^s$ they reduce to a term

$$\mathbb{E}\left[\left(\frac{R(e)}{\tilde{R}(e)}\right)^s \middle| W\right] \leqslant c(s)(q^{1-s} + p^{1+s}) \leqslant c(s)$$

for some suitable $c(s)$ independent of p and q. Putting that together yields

$$\mathbb{E}\left[\prod_{e\in\gamma-e_1}\left(\frac{R(e)}{\tilde{R}(e')}\right)^s\right]=\int_{(0,\infty)^{E(\gamma)}}\mathbb{E}\left[\prod_{e'\in\gamma-e_1}\left(\frac{R(e')}{\tilde{R}(e')}\right)^s\middle|W\right]d\mu(W)\leqslant$$
$$\leqslant\int_{(0,\infty)^{E(\gamma)}}c(s)^{l-1}d\mu(W)=c(s)^{l-1}$$

□

Proof of Lemma 4.2.4. The first problem we have to bring the focus to is the dependence of the edge weights. We are going to construct independent random variables $\bar{R}(e)_{e\in\gamma}$ with $\bar{R}(e)_{e\in\gamma}\mathbb{1}_{D_\gamma}\geqslant\tilde{R}(e)_{e\in\gamma}\mathbb{1}_{D_\gamma}$ that satisfy inequality (4.17). For $e\in\gamma-e_1$ and $k\in\mathbb{N}$ define **independent** Bernoulli variables

$$B_k^e=\text{Bern}\left(\frac{a_e}{k+1+a_e+a_f}\right),\quad B_k^f=\text{Bern}\left(\frac{1+a_f}{2k+1+a_v}\right).$$

f is again the edge preceding e in the path of domination γ and v the vertex incident to e and f. Similarly to (4.15) we define

$$\begin{aligned}
&\bar{N}(e)=\min\{k\geqslant 1:B_k^f=1\}, && \bar{N}(f)=1, && \text{if } B_0^f=0\\
&\bar{N}(f)=\min\{k\geqslant 1:B_k^e=1\}, && \bar{N}(e)=1, && \text{if } B_0^f=1
\end{aligned}$$

and, accordingly,

$$\bar{R}(e)=\frac{\bar{N}(e)}{\bar{N}(f)}.$$

Using this technique we may estimate $\bar{R}(e)$. For $n\geqslant 1$ we have

$$\mathbb{P}(B_0^f=0,\bar{N}(e)=n)=\frac{a_v-a_f}{a_v+1}\frac{1+a_f}{2n+1+a_v}\prod_{k=1}^{n-1}\left(1-\frac{1+a_f}{2k+1+a_v}\right)\leqslant$$
$$\leqslant\frac{a_v}{a_v+1}\frac{1+a_v}{2n}\prod_{k=1}^{n-1}\left(\frac{2k+a_0}{2k+1+a_0}\right)\leqslant\frac{a_0}{2n}K\prod_{k=\lceil a_0/2+1\rceil}^{\lceil a_0/2+n-1\rceil}\frac{2k}{2k+1}\leqslant$$
$$\leqslant\frac{a_0}{2n}K\left(\prod_{k=1}^{\lceil a_0/2\rceil}\frac{2k}{2k+1}\right)^{-1}\prod_{k=1}^{n-1}\frac{2k}{2k+1}\leqslant$$
$$\frac{a_0}{2n}\tilde{C}(K)\prod_{k=1}^{n-1}\exp\left(-\frac{1}{2k}+O\left(k^{-2}\right)\right)\leqslant\tilde{C}(K)a_0n^{-3/2}$$

Observe that, since a_0 is bounded and the term in the middle of the third line is increasing in a_0, we may choose $\tilde{C}$ independent of a_0. Now the condition $s < 1/2$ becomes clear.

$$\begin{aligned}\mathbb{E}\left[\bar{R}(e)^s \mathbb{1}_{B_0^f=0}\right] &\leqslant \sum_{n\geqslant 1} n^s \mathbb{P}\left(B_0^f = 0, \bar{N}(e) = n\right) \leqslant \tilde{C}(K) a_0 \sum_{n\geqslant 1} n^{s-3/2} \\ &\leqslant \tilde{C}(s,K) a_0,\end{aligned}$$

since the sum converges for $s < 1/2$.

$$\mathbb{E}\left[\bar{R}(e)^s \mathbb{1}_{B_0^f=1}\right] \leqslant \sum_{n\geqslant 1} \mathbb{P}(B_n^e = 1) n^{-s} \leqslant \sum_{n\geqslant 1} a n^{-(1+s)} \leqslant \tilde{C}(s) a,$$

since $s > 0$.

Taking $C(s,K) = \tilde{C}(s,K) + \tilde{C}(s)$ proves the claim

$$\mathbb{E}(\bar{R}(e)^s) \leqslant C(s,K) a_0.$$

We still need to show that

$$\bar{R}(e)_{e\in\gamma} \mathbb{1}_{D_\gamma} \geqslant \tilde{R}(e)_{e\in\gamma} \mathbb{1}_{D_\gamma}. \tag{4.18}$$

For this purpose we direct $LRRW(G,a)$ according to the values of B_e^k and B_f^k. Let $v \in V(G)$ be some vertex. If $v \notin \gamma$ nothing needs to be proved. If $v \in \gamma$ and both directed edges e and f have already been traversed from v, $\tilde{R}(e)$ and $\bar{R}(e)$ are already determined and the proof follows from the cases below. Hence, at least one of e and f has not been traversed yet and since D_γ holds, f has been traversed at least once and e has not been traversed yet.

Case 1 v is visited the first time. Let t_1 be the time of the first arrival at v. Since D_γ holds true, this must be along f^-, so e has still routine $w_e(t_1) = a_e$ and f has routine $w_f(t_1) = a_f + 1$. Hence the probability of exiting via f is

$$\frac{1 + a_f}{1 + a_v}.$$

If $B_0^f = 1$, we exit through f (and with suitable probability if $B_0^f = 0$).

Case 2 Later visits to v, $B_0^f = 0$. Let t_n be the time of the n-th visit to v. The current routine of f is (since D_γ holds) at least $w_f(t_n) \geqslant 1 + a_f$ and the current routine of v is $w_v(t_n) = 2n - 1 + a_v$. Hence the probability of exiting via f is

$$\frac{w_f(t_n)}{2n - 1 + a_v} \geqslant \frac{1 + a_f}{2n - 1 + a_v}.$$

If $B_{n-1}^f = 1$ then we exit via f (and with suitable probability also if $B_{n-1}^f = 0$).

Case 3 Later visits to v, $B_0^f = 1$. The edge e has not been traversed and still has initial weight a_e. We are going to ignore transversals of other edges than e and f. So if $LRRW(G,a)$ exits v via one of the edges e and f we direct according to the decisions strategy above. Let t_n be the n-th time v is exited via f. Since f is in the path of domination, the current routine of f is at least $w_f(t_n) = a_f + n$. Hence the probability of exiting via e is

$$\frac{a_e}{w_f(t_n) - a_f + a_v} \leqslant \frac{a_e}{n + a_v} \leqslant \frac{a_e}{n + a_e + a_f}.$$

If $B_{n-1}^e = 0$, we exit via f (and with suitable probability also if $B_{n-1}^e = 1$).

Let us now check that taking the decisions above, $\tilde{R}(e)\mathbb{1}_{D_\gamma} \leqslant \bar{R}(e)\mathbb{1}_{D_\gamma}$ holds. Let $B_0^f = 0$, so either Case 1 or Case 2 needs to be applied. If $n = \min\{k \geqslant 1 : B_k^f = 1\}$ then $LRRW$ has exited v via f by the $(n+1)$-st exit and hence

$$\tilde{R}(e) \leqslant \bar{N}(e) \leqslant n = \bar{R}(e).$$

Let $B_0^f = 1$. Case 3 holds. If $n = \min\{k \geqslant 1 : B_k^e = 1\}$ then f has been used at least n times. Hence

$$\tilde{R}(e) = \frac{1}{\bar{N}(f)} \leqslant \frac{1}{n} = \bar{R}(e).$$

□

We are now ready for the proof of Theorem 4.2.2.

Proof of Theorem 4.2.2. At first, remember that by normalizing $W_{x_0} = 1$, so $W_g = \frac{W_g}{W_{x_0}}$.

$$\begin{aligned}
\mathbb{E}\left[\left(\frac{W_g}{W_{x_0}}\right)^s\right] &\leqslant \mathbb{E}\left[\left(\frac{W_g}{W_{e_1}}\right)^s\right] = \sum_{\gamma \in Y_\Gamma^g} \mathbb{E}\left[\left(\frac{W_g}{W_{e_1}}\right)^s \mathbb{1}_{D_\gamma}\right] \overset{C.S.}{\leqslant} \\
&\leqslant \sum_{\gamma \in Y_\Gamma^g} \mathbb{E}\left[\prod_{e \in \gamma - e_1}\left(\frac{R(e)}{\tilde{R}(e)}\right)^{2s} \mathbb{1}_{D_\gamma}\right]^{1/2} \mathbb{E}\left[\prod_{e \in \gamma - e_1} \bar{R}(e)^{2s} \mathbb{1}_{D_\gamma}\right]^{1/2} \leqslant \\
&\leqslant \sum_{\gamma \in Y_\Gamma^g} \left[C(2s)^{l(\gamma)-1}\right]^{1/2} \left[(C(2s,K)a_0)^{l(\gamma)-1}\right]^{1/2} = \\
&= \sum_{\gamma \in Y_\Gamma^g} (C_0\sqrt{a_0})^{l(\gamma)-1}
\end{aligned} \tag{4.19}$$

by setting $C_0(s,K) = \sqrt{C(2s)C(2s,K)}$. The inequality from the second to the third line follows by lemmas 4.2.3 and 4.2.4. Choosing $s \in (0, 1/4)$ is the best we can do, otherwise Lemma 4.2.4 does not work. Now we choose a_0 sufficiently small so that

$$KC_0\sqrt{a_0} \leqslant \frac{1}{2}. \tag{4.20}$$

This is still not our final a_0, but for the moment it will do. Since the number of simple paths of length l in G is at most K^l and any path to a fixed edge g has length at least $d(g,x_0)+1$ we may bound (4.19) by

$$\sum_{\gamma\in Y_\Gamma^g}(C_0\sqrt{a_0})^{l(\gamma)-1} \leqslant \sum_{l\geqslant d(g,x_0)+1} K^l(C_0\sqrt{a_0})^{l-1} \tag{4.21}$$

$$= \sum_{l\geqslant d(g,x_0)} K(KC_0\sqrt{a_0})^l \leqslant 2K(KC_0\sqrt{a_0})^{d(g,x_0)}. \tag{4.22}$$

Redefining $C = C_0K$ proves Theorem 4.2.2. □

4.2.2 Tight Measures, Borel-Cantelli and Markov's Inequality

Proof of Theorem 4.2.1. Let $G_n = G[\{v \in V(G) : d(v,G) \leqslant n\}]$. Again we denote the mixing measure for G_n by $\mu^{(n)}$. Denote the weight for an edge g in G_n by W_g^n. Recall from Theorem 3.6.11 that the measures $\mu^{(n)}$ are a tight on $(0,\infty)^{E(G_n)}$ and thus there exists a convergent subsequence with limit μ^*. Now let $s \in (0,1/4)$. Note that for any edge $g \in E(G_n)$ we have by Markov's inequality and Theorem 4.2.2

$$\mu^{(n)}(W_g^n > w) = \mu^{(n)}\left((W_e^n)^s > w^s\right) \leqslant \frac{2K(C\sqrt{a_0})^{d(g,x_0)}}{w^s}.$$

This inequality holds for all $\mu^{(n)}$ and hence also for any weak limit μ^*. Take $w = (2K)^{-d(g,x_0)}$. Observe that the number of edges at distance l is at most K^{l+1}. Hence the probability of having an edge g at distance $l = d(g,x_0)$ with weight $W_g > w$ can be bounded by

$$\begin{aligned}\mathbb{P}(\exists e \in E(G) : d(e,x_0) = l, W_e > w) &\leqslant K^{l+1}(2K)^{sl}2K(C\sqrt{a_0})^l = \\ &= 2K^2(2^sK^{1+s}c\sqrt{a_0})^l. \end{aligned} \tag{4.23}$$

Now re-choose a_0 so that $2^sK^{1+s}C\sqrt{a_0} \leqslant \frac{1}{2}$. This is our final a_0.
(4.23) may be bounded by

$$K^2 2^{1-l}.$$

The sequence $K^2 2^{1-l}$ is summable. Thus by the Borel-Cantelli lemma μ-almost surely the number of edges e violating $W_e \leqslant (2K)^{-d(e,x_0)}$ is finite. The total weight of the edges is almost surely a finite number and hence a stationary measure $\pi_{\mathbb{W}}$ exists almost surely, given explicitly in terms of a realization of the weights $\mathbb{W}$ by

$$(\pi_{\mathbb{W}}(v))_{v\in V(G)} = \left(\frac{\mathbb{W}_v}{\sum_{v\in V(G)} \mathbb{W}_v}\right)_{v\in V(G)}.$$

□

Chapter 5

Properties of the Mixing Measure

5.1 Initial Vertices and Measures

Needless to say, the mixing measure depends on the initial vertex. It would be nice to find an explicit expression, i.e. a density w.r.t. some invariant measure. In the finite case this has been done by Merkl, Öry and Rolles, see [14]. However, for infinite trees the mixing measure is in general not absolutely continuous w.r.t. Lebesgue measure on a cartesian product of simplices. This follows directly by arguing that the product of infinitely many Dirichlet densities is, except for some special cases, not a bounded function. But it is still reasonable to compare the mixing measures for different initial vertices. At first we will focus on two adjacent initial vertices x_0 and x_1. Let a_0 and a_1 be the total initial weight of edges incident to x_0 and x_1, respectively. For trees the identity

$$d\mu_0(P) = \frac{\Gamma\left(\frac{a_0}{2}\right)\Gamma\left(\frac{a_1+1}{2}\right)}{\Gamma\left(\frac{a_1}{2}\right)\Gamma\left(\frac{a_0+1}{2}\right)}\sqrt{\frac{p_{1,0}}{p_{0,1}}}d\mu_1(P) \tag{5.1}$$

is not too difficult to prove, thinking about each vertex and the star graph having v as its center as an urn. In the finite case the mixing measure has a density w.r.t. Lebesgue measure, given explicitly by the product of some Dirichlet densities. Thus to show (5.1) all factors but one cancel out, passing to the limit (5.1) holds still true. For general graphs we need a different approach. The following result has been obtained in the course of the elaboration of this thesis.

Theorem 5.1.1. *Let (G, a, x_0) be a recurrent instance for LRRW with initial vertex x_0, (G, a, x_1) the same instance but starting from x_1. Denote their mixing measures by μ_0 and μ_1, respectively. μ_0 and μ_1 are mutually absolutely continuous and their derivative is given by* (5.1).

Clearly, if (G, a, x_0) is recurrent then (G, a, x_1) needs to be recurrent, too. To prove the theorem we will make use of paths starting and ending in x_0 and x_1, respectively. Recall that, if G is finite, x_0 as well as x_1 are visited infinitely often. Interestingly, the derivative (5.1) depends only on the total initial weights a_0 and a_1 and the values of $p_{0,1}$ and $p_{1,0}$ but not on any other values. We first start by expressing the probability of a certain set of stochastic matrices by a set of finite paths. The following lemma is mainly important to introduce some new notation, the convergence result may as well be obtained using that $LRRW$ is a mixture of Markov chains. In the following it is assumed that $LRRW(G, a)$ is recurrent, let x^* be some vertex, in the following called reference vertex. x^* need not be necessarily the initial one. Denote by μ_0 the unique mixing measure for initial vertex x_0. Let $e_1 = (v_1, w_1), \ldots, e_k = (v_k, w_k)$ be a finite set of directed edges in the directed version of G. Let $I_1, \ldots, I_k \subset [0,1]$ be open nonempty intervals. We denote by $\mathcal{P}'$ the set of reversible stochastic matrices $P = p(\cdot,\cdot)$ on $V(G) \times V(G)$ that satisfy

$$p(e_i) \in I_i, i = 1, \ldots, k. \tag{5.2}$$

For a finite path $y = (u_0, \ldots, u_l)$, a vertex v and an edge $e = \{v, w\}$ define

$$N_v(y) = |\{1 \leqslant t \leqslant l : u_t = v\}|, \tag{5.3}$$
$$N_e(y) = |\{0 \leqslant t \leqslant l-1 : \{u_t, u_{t+1}\} = \{v, w\} \vee \{u_t, u_{t+1}\} = \{w, v\}\}|. \tag{5.4}$$

Observe that the first appearence of x_0 does not count for $N_v(x_0)$. This is not essential for the statement or the proof of the following two lemmas but makes things slightly easier.

Lemma 5.1.2. *Let $Y_n := Y_n^{x_0,x_0}(\mathcal{P}')$ be the set of finite paths y that satisfy the following.*

1. *$y \in Y_n$ starts and ends in x_0.*
2. *$N_{x_0}(y) = n$.*
3. *For $e_i = (v_i, w_i)$, $N_{v_i}(y) =: N_i^v$ and $N_{e_i}(y) =: N_i^e$ satisfy*

$$\frac{N_i^e}{N_i^v} \in I_i.$$

We call Y_n the approximating set of $\mathcal{P}'$ for initial vertex x_0 and reference vertex x_0. Then

$$\mu_0(\mathcal{P}') = \lim_{n\to\infty} \mathbb{P}(Y_n).$$

Proof. The random variables $\frac{N_i^e}{N_i^v}$ converge for $n \to \infty$ to the same random variable as

$$\frac{w_{e_i}(t)}{w_{v_i}(t)} = \frac{N(e_i,t) + a_e}{\sum_{w\in\mathcal{N}_G(v_i)} N(\{v,w\},t) + a_{v,w}}$$

does for $t \to \infty$, their convergence being guaranteed since recurrent *LRRW* is a unique mixture of Markov chains. The dependence on N_{x_0} has no impact since x_0 is visited infinitely often almost surely anyway. Of course, the choice of the reference vertex does not have any impact either. Since the random variables $\frac{N_i^e}{N_i^v}$ may be expressed as a function of $y \in Y_n$ the claim of the lemma holds. □

Now that we may approximate the mixing measure by the probability of a certain set of paths we prove Theorem 5.1.1.

Proof of Theorem 5.1.1. Let (G, a, x_0) be a recurrent instance of *LRRW* with initial vertex x_0, (G, a, x_1) the same instance but with initial vertex x_1. Clearly, changing the initial vertex does not affect recurrence. Let $e_{0,1} = (x_0, x_1), e_{1,0} = (x_1, x_0), e_1, \dots, e_k$ be a set of directed edges. Let $0 < p_{0,1} := p(x_0, x_1), p_{1,0} := p(x_1, x_0), p_1 := p(e_1), \dots, p_k = p(e_k) < 1$. We assume that these p admit a reversible Markov chain, otherwise equation (5.1) holds trivially true. We collect the p's in a vector $p \in (0,1)^{k+2}$, let $\varepsilon \in (0,1)^{k+2}$ so that

$$0 < p - \varepsilon < p + \varepsilon < 1 \tag{5.5}$$

holds componentwise. Denote by $\mathcal{P}^\varepsilon$ the set of stochastic matrices $P' = p'(\cdot, \cdot)$ on $V(G) \times V(G)$ associated with reversible chains that satisfy

$$p - \varepsilon \leqslant p' \leqslant p + \varepsilon.$$

For simplicity we will assume that $\varepsilon = \epsilon \mathbb{1}^{k+2}$, where $\mathbb{1}^{k+2}$ is the all-ones vector in $\mathbb{R}^{k+2}$. This does not affect any step of the proof. Take $a_0 := a_{x_0}$, $a_1 := a_{x_1}, a_{0,1} = a_{x_0,x_1}$. Denote by $\mathbb{P}_0$ and $\mathbb{P}_1$ the probability measures of $LRRW(G, a)$ for initial vertices x_0 and x_1, respectively. Let $Y_n^0 := Y_n^{x_0,x_0}(\mathcal{P}^\varepsilon)$ and $Y_n^1 := Y_n^{x_0,x_1}$ be the approximating sets for $\mathcal{P}^\varepsilon$ and initial vertices x_0 and x_1, respectively. The reference vertex for both approximating sets is x_0. Clearly,

$$\lim_{n \to \infty} \frac{P_0(Y_n^0(\mathcal{P}^\varepsilon))}{P_1(Y_n^1(\mathcal{P}^\varepsilon))} = \frac{\mu_0(\mathcal{P}^\varepsilon)}{\mu_1(\mathcal{P}^\varepsilon)}.$$

We would like to estimate of the ratio $\frac{\mathbb{P}(Y_n^0)}{\mathbb{P}_1(Y_n^1)f}$. Let $\mathfrak{S}_l$ be the set of cyclic permutations $(\rho_j)_{j=0,\dots,l}$ on $\{0, \dots, l\}$ of the form

$$\rho_j(i) = \begin{cases} i + j - (l+1) & \text{if } i + j \geqslant l + 1 \\ i + j & \text{else} \end{cases}.$$

We now regard $y = (v_0, \dots, v_{l+1}) \in Y_n^0$ as a cyclic object. We thus omit to write v_{l+1}, being equal to v_0 anyway. Consider the equivalence classes

$$\begin{aligned} [y] &= \{(v_{\rho_j(i)})_{i=0,\dots l} : j \in \{0, \dots, l\}\} \\ [y]_0 &= \{(v_{\rho_j(i)})_{i=0,\dots l} : j \in \{0, \dots, l\}, v_j = x_0\} \\ [y]_1 &= \{(v_{\rho_j(i)})_{i=0,\dots,l} : j \in \{0, \dots, l\}, v_j = x_1\}. \end{aligned}$$

It is not difficult to see that $\frac{|[y]_0|}{|[y]_1|} = \frac{N_{x_0}(y)}{N_{x_1}(y)}$, so that

$$\frac{\mathbb{P}_0([y]_0)}{\mathbb{P}_1([y]_1)} = \frac{N_{x_0}(y)}{N_{x_1}(y_1)} \frac{\mathbb{P}_0(y_0)}{\mathbb{P}_1(y_1)},$$

where y_0 and y_1 are representatives of $[y]_0$ and $[y]_1$, respectively. Note that

$$\frac{p_{1,0}-\epsilon}{p_{0,1}+\epsilon} \leqslant \frac{N_{x_0}(y)}{N_{x_1}(y)} \leqslant \frac{p_{1,0}+\epsilon}{p_{0,1}-\epsilon} \tag{5.6}$$

But for all $y \in Y_n^0$, $[y]_0 \subset Y_n^0$ and $[y]_1 \subset Y_n^1$ holds true. Splitting up Y_n^0 and Y_n^1 into classes w.r.t. $[\cdot]_0$ and $[\cdot]_1$, respectively, yields

$$\frac{p_{1,0}-\epsilon}{p_{0,1}+\epsilon} \inf\left\{\frac{\mathbb{P}_0(y_0)}{\mathbb{P}_1(y_1)}, y_0 \in Y_n^0, y_1 \in Y_n^1, [y_0] = [y_1]\right\} \leqslant \frac{\mathbb{P}_0(Y_n^0)}{\mathbb{P}_1(Y_n^1)} \leqslant$$
$$\leqslant \frac{p_{1,0}+\epsilon}{p_{0,1}-\epsilon} \sup\left\{\frac{\mathbb{P}_0(y_0)}{\mathbb{P}_1(y_1)}, y_0 \in Y_n^0, y_1 \in Y_n^1, [y_0] = [y_1]\right\}.$$

The expression $\frac{\mathbb{P}_0(y_0)}{\mathbb{P}_1(y_1)}$ depends only on $N_{x_0}(y_0) =: n$ and $N_{x_1}(y_0) =: N_1$. We only deal with the case $n \geqslant N_1$, the other case works the same way. Straightforward calculation yields

$$\frac{P_0(y_0)}{P_1(y_1)} = \underbrace{\prod_{\substack{i=0 \\ i \text{ even}}}^{2(N_1-1)} \frac{(a_1+i)(a_0+i+1)}{(a_0+i)(a_1+i+1)}}_{=:L_1(N_1)} \underbrace{\prod_{\substack{i=2N_1(y)-1 \\ i \text{ even}}}^{2(n-1)} \frac{a_0+i+1}{a_0+i}}_{=:L_2(\frac{N_1}{n},n)} =: L(N_1,n).$$

Since $L(N_1,n)$ is decreasing in N_1 and the ratio $\frac{N_0}{N_1}$ may be estimated using (5.6) we choose

$$N_1^+(\epsilon,n) := n\left\lceil\frac{p_{1,0}+\epsilon}{p_{0,1}-\epsilon}\right\rceil, N_1^-(\epsilon,n) := n\left\lfloor\frac{p_{1,0}-\epsilon}{p_{0,1}+\epsilon}\right\rfloor.$$

Hence

$$L(N_1^+(\epsilon,n)) \leqslant \frac{P_0(y_0)}{P_1(y_1)} \leqslant L(N_1^-(\epsilon,n)).$$

To determine the value of L we use the representation of the Gamma function

$$\Gamma(x) = \lim_{n\to\infty} \frac{n!n^x}{x(x+1)\cdots(x+n)}.$$

$L_1(N_1)$ converges to

$$\frac{\Gamma(\frac{a_0}{2})\Gamma(\frac{a_1+1}{2})}{\Gamma(\frac{a_1+1}{2})\Gamma(\frac{a_0}{2})} =: \Gamma_{0,1}$$

and $L_2(x,n)$ converges pointwise to $\sqrt{\frac{1}{x}}$.

So, letting $n \to \infty$ we obtain

$$\frac{p_{1,0}-\epsilon}{p_{0,1}+\epsilon}\Gamma_{0,1}\sqrt{\frac{p_{0,1}-\epsilon}{p_{1,0}+\epsilon}} \leqslant \frac{\mu_0(P^\epsilon)}{\mu_1(P^\epsilon)} \leqslant \frac{p_{1,0}+\epsilon}{p_{0,1}-\epsilon}\Gamma_{0,1}\sqrt{\frac{p_{0,1}+\epsilon}{p_{1,0}-\epsilon}}.$$

Letting $\epsilon \to 0$ and observing that sets of the type $\mathcal{P}^\epsilon$ generate $\mathcal{B}(\mathcal{P})$ finishes the proof. $\square$

For shortness we will now write

$$\Gamma_{v,w} := \frac{\Gamma\left(\frac{a_v}{2}\right)\Gamma\left(\frac{a_w+1}{2}\right)}{\Gamma\left(\frac{a_w}{2}\right)\Gamma\left(\frac{a_v+1}{2}\right)},$$

so for two adjacent vertices v, w

$$d\mu_v(P) = \Gamma_{v,w}\sqrt{\frac{p(w,v)}{p(v,w)}}d\mu_w(P).$$

Write $d_v = \deg_G(v)$. For all $t \in \mathbb{R}^{d_v}, \sum_{w\in\mathcal{N}_G(v)} t_w = 1$

$$d\mu_v(P) = \sum_{w\in\mathcal{N}_G(v)} t_w\Gamma_{v,w}\sqrt{\frac{p(w,v)}{p(v,w)}}d\mu_w(P). \tag{5.7}$$

This holds especially for $t_w = p(v,w)$ and hence for all v

$$d\mu_v(P) = \sum_{w\in\mathcal{N}_G(v)} \Gamma_{v,w}\sqrt{p(w,v)p(v,w)}d\mu_w(P). \tag{5.8}$$

Setting $d\mu = (d\mu_v)_{v\in V(G)}, P_\Gamma = \left(\Gamma_{v,w}\sqrt{p(v,w)p(w,v)}\right)_{v,w\in V(G)}$ we may write equation (5.8) as

$$d\mu(P) = P_\Gamma d\mu(P) \Leftrightarrow (P_\Gamma - 1)d\mu(P) = 0. \tag{5.9}$$

Now let μ_v^* be a mixing measure for initial vertex v, obtained as a weak limit of a sequence of measures $(\mu_{v,n_k})_{k\in\mathbb{N}}$ on $(G^{n_k})_{k\in\mathbb{N}}$. Let $(\mu_w^*)_{w\in\mathcal{N}(v)}$ be weak limits of (μ_{w,n_k}). Let $e_1 = (v_1,w_1),\ldots,e_k = (v_k,w_k)$ be a finite set of directed edges, let $I_1\ldots,I_k \in \mathcal{B}([0,1])$. If $\mathcal{P}' = (p'(v,w))_{v,w\in V(G)} \in \mathcal{B}(\mathcal{P})$ is a set of the form

$$\{P' \in \mathcal{P} : p'(e_i) \in I_i\}$$

then equation (5.8) still holds, the functions $\Gamma_{v,w}\sqrt{p(v,w)p(w,v)}$ being continuous.

5.2 Uniqueness of the Mixing Measure on a Sub-σ-algebra

We would like to prove uniqueness of the mixing measure in the general case. A way of doing so is to use Lévy's Continuity Theorem. But to do so we need finite moments of the respective random variables. For this purpose the notation of μ as a measure on the space of stochastic matrices is more convenient. The following lemma and its proof have been obtained in the course of the elaboration of this thesis.

Lemma 5.2.1. *Let E_n be a finite set of edges in G, $|E_n| = n$. For an edge $e = \{v, w\} \in E_n$ denote by $p(e) := p(v, w)p(w, v)$ the product of the transition probabilities in either direction. The joint distribution of $P_n := (p(e))_{e\in E_n}$ is uniquely determined by G and a, thus independent of the choice of the mixing measure.*

We first give a general idea of the proof strategy. Let x_0 be the initial vertex, let x_1 be adjacent to x_0 in G. We show uniqueness of the distribution of $p := p(x_0, x_1)p(x_1, x_0)$. By Levy's continuity theorem the distribution of p is uniquely determined by its moment generating function $F_p(t) = \mathbb{E}[e^{tp}]$, provided that it is bounded in a neighborhood of 0. But p is bounded by 1, so $F_p(t)$ is finite on all $\mathbb{R}$. The moments p^k are determined by (G, a) since

$$\int_{\mathcal{P}} p^k d\mu^*$$

is exactly the probability of moving k times forth and back on $\{x_0, x_1\}$. This is simply an event which may be expressed in terms of a. Thus the distribution of p is uniquely determined.

Proof. If μ^* is a mixing measure, it induces a probability measure on P_n. W.l.o.g. we assume that E_n is connected and contains an edge incident to x_0, otherwise we make E_n bigger by adding some edges. We show that all moments $\prod_{e\in E_n} p(e)^{k_e}$ are determined. Then the moment generating function

$$M_{P_n} : \mathbb{R}^n \to \mathbb{R}, \quad M_{P_n}(t) = \mathbb{E}[e^{t^\top P_n}]$$

is determined and obviously bounded on all $\mathbb{R}^n$. Lévy's continuity theorem then implies uniqueness of the distribution. Now let y_n be a finite path that starts and ends in x_0 and traverses all edges in E_n at least once in both directions. Let $\tilde{y}_n$ be some path that traverses each edge e additionally k_e times back and forth. Denote by $p_n := \prod_{e\in E(y_n)} p(e)$ and $\tilde{p}_n = p_n \prod_{e\in E_n} p(e)^{k_e}$ the probabilities of y_n and $\tilde{y}_n$, respectively, w.r.t. some Markov chain. Denote by y^m and $\tilde{y}^m$ the concatenation of m versions of y and $\tilde{y}$, respectively. Observe that for all l_1, l_2 the expression

$$\int_{\mathcal{P}} p_n^{l_1} \tilde{p}_n^{l_1} d\mu^* = \mathbb{P}([y^{l_1}, \tilde{y}^{l_2}])$$

is in fact the probability of a certain path, uniquely determined by G and a and thus independent of the choice of the mixing measure. By Lévy's continuity theorem the joint distribution of $(p_n, \tilde{p}_n)$ is unique. But hence the distribution of

$$\prod_{e \in E_n} p(e)^{k_e} = \frac{\tilde{p}_n}{p_n} \tag{5.10}$$

is unique and so its expectation is determined. Since all moments are of the form in (5.10), M_{P_n} is determined which proves the claim of the lemma. □

Corollary 5.2.2. *The distribution of the matrix P_Γ from the last lemma is unique.*

Proof. P_Γ is a function of $p(e)_{e \in E(G)}$. □

The last two chapters give reason to believe that for any locally finite graph G and for any initial weights a the mixing measure for $LRRW(G, a)$ is unique.

Chapter 6

Tables

Table 6.1: Numerical approximations of the critical values $a_0(K+1)$ of phase transition for $LRRW$ on $(K+1)$-regular trees for $K = 0, \ldots, 14$.

K	$a_0(K)$
1	∞
2	0.232910211931729
3	0.123919276013275
4	0.0847016129169972
5	0.0643922303308135
6	0.0519539698295227
7	0.0435487942006840
8	0.0374870609401862
9	0.0329079218397183
10	0.0293264016927345
11	0.0264483501283220
12	0.0240849558006896
13	0.0221094513992165
14	0.0204335560602396
15	0.0189938964084764

Table 6.2: List of symbols

Symbol	Meaning
$\mathbb{N}$	natural numbers including 0
$\mathbb{N}^*$	natural numbers without 0
$\mathbb{Z}$	integer numbers
$\mathbb{Q}$	rational numbers
$\mathbb{R}$	real numbers
Z	some countable state space
$\mathfrak{S}_n$	symmetric group on a set with cardinality n
$G = (V, E)$	a locally finite graph with vertex set V, edge set E
$G[V']$	the graph induced by a subset V' of the original vertices
$T = (V, E)$	a locally finite tree
$S = (V, E)$	a finite star graph
$\deg_G(v)$	number of vertices adjacent to vertex v in G
K	maximum degree in a graph or number of children in case of a tree
$\mathcal{N}_G(v)$	set of vertices adjacent to a vertex v in G
$\delta_G(v)$	set of edges incident to a vertex v in G
$\delta_G^+(v)$	outgoing edges in a directed graph
$d(\cdot, \cdot)$	shortest path metric
X	stochastic process on Z
$(\mathcal{X}, \tau)$	a Polish space
$\mathcal{B}(\cdot)$	σ-algebra on a set
$\mathcal{E}$	exchangeable σ-algebra
x_0	initial vertex for X, $X_0 = x_0$ almost surely
y	a path in G
y^+, y^-	a directed path and its reversal
Y	random variable on the space of x_0-x_0-paths or general $\mathcal{X}$-valued random variable
$\mathcal{Y}$	set of finite paths
γ	path of domination
$\mathbb{P}$	probability measure w.r.t. some process
$\mathbb{E}[\cdot]$	expectation of a random variable or distribution
$\mathrm{Var}[\cdot]$	variance of a random variable or distribution
$\mathcal{F}, \mathcal{G}$	filtrations
M	martingales and backwards martingales
$LRRW$	linearly reinforced random walk
B_n	chosen ball in Pólya urn process at time n
U_n	relative content in Pólya's urn process at time n

Continued on next page

Table 6.2 – continued from previous page

Symbol	Meaning
$N(e,t)$	number of transversals of edge e up to time t
a	initial edge weights $(a_e)_{e\in E(G)}$
$w_e(t)$	routine of edge e at time t, $w_e(t) = a_e + N(e,t)$ for *LRRW*
W_e	the weight of an edge, seen as a random variable
$\tilde{W}_e$	like W_e, but normalized w.r.t. some constant
μ	distribution or operator-valued probability measure
P	transition matrix of a reversible chain
$\mathcal{P}$	set of probability measures or
	set of reversible stochastic matrices
$\mathbb{1}_x$	the point measure at x
$\mathbb{1}_A$	the indicator function on the set A
$\lhd$	convex order
$\mathbb{W}$	fixed weights on the edges of a graph
$\Gamma(\cdot)$	the Gamma function
$\beta(\cdot,\cdot)$	the Beta function
κ, C, c	functions continuous in all arguments serving to bound expressions

Bibliography

[1] Omer Angel, Nicholas Crawford, and Gady Kozma. "Localization for Linearly Edge Reinforced Random Walks". Version 1. In: *Duke Math* 5 (Mar. 19, 2012), pp. 889–921. URL: http://arxiv.org/abs/1203.4010v1.

[2] Herman Chernoff. "A Measure of Asymptotic Efficiency for Tests of a Hypothesis Based on the sum of Observations". In: *Ann. Math. Statist.* 23.4 (Dec. 1952), pp. 493–507. DOI: 10.1214/aoms/1177729330. URL: https://doi.org/10.1214/aoms/1177729330.

[3] Edward A. Codling, Michael J. Plank, and Simon Benhamou. "Random walk models in biology". In: *J R Soc Interface* 5.25 (2008), pp. 813–834. ISSN: 1742-5689. DOI: 10.1098/rsif.2008.0014. URL: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2504494/.

[4] Djurdje Cvijović. "New integral representations of the polylogarithm function". In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* (Dec. 2006), pp. 897–905. DOI: 10.1098/rspa.2006.1794. URL: https://doi.org/10.1098/rspa.2006.1794.

[5] Burgess Davis. "Reinforced random walk". In: *Probability Theory and Related Fields* (June 1, 1990), pp. 203–229. DOI: 10.1007/BF01197845. URL: https://doi.org/10.1007/BF01197845.

[6] P. Diaconis and D. Freedman. "De Finetti's Theorem for Markov Chains". In: *Ann. Probab.* 8.1 (Feb. 1980), pp. 115–130. DOI: 10.1214/aop/1176994828. URL: http://dx.doi.org/10.1214/aop/1176994828.

[7] T. E. Harris. *The Theory of Branching Processes*. Die Grundlehren der Mathematischen Wissenschaften in Einzeldarstellungen. Springer-Verlag, 1963. URL: https://books.google.at/books?id=%5C_ORQAAAAMAAJ.

[8] E. Hewitt and L. J. Savage. "Symmetric measures on Cartesian products". In: *Trans. Amer. Math. Soc* 80 (Nov. 1955), pp. 470–501.

[9] Olav Kallenberg. *Foundations of modern probability*. Second. Probability and its Applications (New York). Springer-Verlag, New York, 2002, pp. 156–169. ISBN: 0-387-95313-2. DOI: 10.1007/978-1-4757-4015-8. URL: http://dx.doi.org/10.1007/978-1-4757-4015-8.

[10] Daniel Kious and Vladas Sidoravicius. "Phase transition for the Once-reinforced random walk on $\mathbb{Z}^d$-like trees". Version 1. In: *arXiv* (Apr. 26, 2016). URL: http://arxiv.org/abs/1604.07631v1.

[11] Gady Kozma. "Reinforced random walk". In: *arXiv* (2012). URL: https://arxiv.org/abs/1208.0364.

[12] Howard A. Levine et al. "Mathematical modeling of capillary formation and development in tumor angiogenesis: Penetration into the stroma". In: *Bulletin of Mathematical Biology* 63.5 (2001), pp. 801–863.

[13] Vlada Limic. "Attracting edge property for a class of reinforced random walks". In: *Ann. Probab.* 31.3 (July 2003), pp. 1615–1654. DOI: 10.1214/aop/1055425792. URL: http://dx.doi.org/10.1214/aop/1055425792.

[14] Franz Merkl, Aniko Öry, and Silke W. W. Rolles. "The 'magic formula' for linearly edge-reinforced random walks". In: *Statistica Neerlandica* 62.3 (2008), pp. 345–363. ISSN: 1467-9574. DOI: 10.1111/j.1467-9574.2008.00402.x. URL: http://dx.doi.org/10.1111/j.1467-9574.2008.00402.x.

[15] Franz Merkl and Silke W. W. Rolles. "Linearly edge-reinforced random walks". In: *Dynamics & Stochastics*. Ed. by Dee Denteneer, Frank den Hollander, and Evgeny Verbitskiy. Vol. Volume 48. Lecture Notes–Monograph Series. Beachwood, Ohio, USA: Institute of Mathematical Statistics, 2006, pp. 66–77. DOI: 10.1214/074921706000000103. URL: http://dx.doi.org/10.1214/074921706000000103.

[16] Robin Pemantle. "A survey of random processes with reinforcement". In: *Probab. Surveys* 4 (2007), pp. 1–79. DOI: 10.1214/07-PS094. URL: http://dx.doi.org/10.1214/07-PS094.

[17] Robin Pemantle. "Phase Transition in Reinforced Random Walk and RWRE on Trees". In: *Ann. Probab.* 16.3 (July 1988), pp. 1229–1241. DOI: 10.1214/aop/1176991687. URL: http://dx.doi.org/10.1214/aop/1176991687.

[18] L. C. G. Rogers and David Williams. "Diffusions, Markov Processes, and Martingales, Vol. 2, Itô Calculus". In: *Applied Stochastic Models and Data Analysis* 4.2 (1988). ISSN: 1099-0747. DOI: 10.1002/asm.3150040208. URL: http://dx.doi.org/10.1002/asm.3150040208.

[19] Silke W. W. Rolles. "How edge-reinforced random walk arises naturally". In: *Probability Theory and Related Fields* 126.2 (2003), pp. 243–260.

[20] Christophe Sabot and Pierre Tarrès. "Edge-reinforced random walk, Vertex-Reinforced Jump Process and the supersymmetric hyperbolic sigma model". In: *arXiv* (2011). URL: `https://arxiv.org/pdf/1111.3991.pdf`.

[21] Stanford University. Department of Statistics, P. Diaconis, and National Science Foundation (U.S.) *Recent Progress on de Finetti's Notions of Exchangeability*. 1988. URL: `https://books.google.at/books?id=Gw3sHgAACAAJ`.