Dipl.-Ing. Wilfried Huss

# Internal aggregation models

## DISSERTATION

zur Erlangung des akademischen Grades einer/s Doktorin/Doktors
der technischen Wissenschaften/Naturwissenschaften

Graz University of Technology

Technische Universität Graz

Betreuer/in:
Univ.-Prof. Dipl.-Ing. Dr.techn. Wolfgang Woess

Institut für Mathematische Strukturtheorie

Graz, im Oktober 2010

## Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitely marked all material which has been quotes either literally or by content from the used sources.

. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

date (signature)

# Contents

## Chapter 1

# Introduction

The purpose of this thesis is the study of several related growth models on non homogeneous state spaces. The growth models in question are *Internal Diffusion Limited Aggregation* (IDLA), *Rotor-Router Aggregation* and also the so-called *Divisible Sandpile Model*. All three models share many properties even though IDLA is a probabilistic growth model based on random walks, while the other two are strictly deterministic. In fact the divisible sandpile model has been introduced by LEVINE AND PERES [LP09] as a tool to prove shape theorems for IDLA and the Rotor Router model.

In both internal diffusion limited aggregation, which is a special case of a process which DIACONIS AND FULTON introduced in [DF91], and Rotor Router aggregation particles move on the vertices of a graph, until they reach a site which is unoccupied. There the particle stops and from now on occupies this site, and a new particle starts its journey at the origin. The main point of research in this area is how the set of occupied sites behaves – for example, if it has a limiting shape, if properly rescaled. The difference between the two growth models lies in the rule which governs the movement of the particles. In IDLA each particle performs a random walk, where the law of each particle is the same and all random walks are independent of each other. In the rotor-router model, on the other hand, the particles perform deterministic walks, where each particle follows an *arrow* (or *rotor*, hence the name of the model) which points to one of the neighbours of the current site. But before the particle moves, the arrow is changed to point to the next of the neighbours, in a previously chosen order. When a particle stops and the next one starts at the origin, it is crucial that the arrows are not reset, but remain in the same state. Hence, in contrast to IDLA, in the rotor-router model the particles are in fact dependent on all the previous particles.

The third growth model, the divisible sandpile, differs in that here not individual particles are moving around, but a distribution of mass, where each vertex can have an arbitrary amount of mass, which changes in time. At each timestep one vertex distributes mass to all its neighbours and keeps only a fixed amount of mass for itself. Since the mass is continuous, the process never terminates and one needs to consider the limit of the process. One is mainly interested in the set of vertices where the limit mass distribution is nonzero.

It turns out that, although all three models look quite different, their behaviour is very similar. This was first noticed by Levine and Peres [LP09, LP10] for the case when the state space is an Euclidean lattice. Computer simulations suggest that the connection between the three growth models holds in wide generality, but only partial results are available for other state spaces.

What makes it possible to prove theoretic results, is the so called *Abelian property*, which all three growth models possess. In the case of IDLA and the rotor-router model this means that, if we let several particles run at the same time, instead of one after another, it is irrelevant for the end result in which order the particles make their moves. In the case of the divisible sandpile model, the Abelian property says that the order in which vertices distribute their excess mass does not affect the limiting distribution. Because of this property our growth models also have a close connection with the so called *Abelian sandpile model* and the *Abelian sandpile group* – see [HLM$^+$08] or [MRZ01] for an introduction. But we will not explore this connection in detail.

## 1.1 Overview

Chapter 2 starts with an introduction to Markov chains, especially random walks on graphs, which allows us to fix the notation we will use in the rest of this thesis. Next, we will give the definitions of the three growth models under consideration, and finally the chapter concludes with a brief overview on the available literature in this area.

In Chapter 3 we will prove the Abelian property and the existence of a limiting mass distribution for the divisible sandpile, which is associated to a reversible random walk on an arbitrary graph. Morevover, we give a description of its limiting shape in terms of a discrete obstacle problem. This an extension of a result of Levine and Peres [LP09]. We will then apply this result to the divisible sandpile on the comb, that is, the spanning tree of $\mathbb{Z}^2$ which is constructed by deleting all horizontal edges except the ones on the $x$-axis. In Section 3.5 we obtain an explicit description of the limiting shape of the divisible sandpile in that case.

In Chapter 4 we will derive a shape theorem for the rotor-router model on the comb, using the results of Chapter 3, and applying an idea of HOLROYD AND PROPP [HP10]. They define a weight function on the space of all configurations of the rotor-router model, which is invariant under the movement of particles. Using this method, we are only able to give a relatively weak inner bound for the rotor-router model, which holds for all possible initial states of the rotors, and all rotor sequences. On the other hand, if we fix a special initial rotor configuration we can give the exact shape of the rotor-router cluster, using an idea of KAGER AND LEVINE [KL]. Section 4.4 gives a general inner bound for the rotor-router cluster, in terms of the divisible sandpile cluster of the same graph. This result holds for all regular graphs, but provides only a relativly weak inner bound.

In Chapter 5 we give an application of the rotor-router model for calculating the harmonic measure of finite subsets of the comb. This again makes use of the rotor weights we introduced in Chapter 4. In particulary we will calculate the harmonic measure of the rotor-router cluster as it will be useful in the study of IDLA on the comb. Since we will make use of singularity analysis for linear differential equations, we give a short overview of this theory in Section 5.1.

Chapter 6 is devoted to Internal Diffusion Limited aggregation on the comb. We will prove a shape theorem using the results of Chapter 3, for the inner bound. The proof is based on ideas of LAWLER, BRAMSON AND GRIFFEATH [LBG92] and LEVINE AND PERES [LP09].

In Chapter 7 we examine again IDLA, this time on non-amenable graphs, that is, graphs where the random walk which governs the particles has a spectral radius which is strictly smaller than one. We will give a shape theorem based on level sets of the Green function of the random walks. This is an extension of a shape theorem of BLACHÈRE AND BROFFERIO [BB07] for IDLA on finitely generated groups with exponential growth. The results of this chapter have been previously published in [Hus08].

Finally, in Chapter 8 we give an outlook to future work and some open problems, and conclude with some simulation results for IDLA and rotor-router aggregation on the graphical Sierpinski carpet, where no theoretical results are yet available.

Chapters 4, 5 and 6 are based in part on joint work with Ecaterina Sava.

## Chapter 2

# Preliminaries

In this chapter we first introduce random walks on graphs and some basic facts and tools related to random walks theory, which will be needed later on. Then we will give precise definitions of the three growth models, and give a short overview of the available literature in this area. We conclude each section with the main theorems, which will be proved in the later chapters.

## 2.1 Random Walks

Let $G = (V, E)$ be an infinite, locally finite, connected graph with vertex set $V$ and edge set $E$. When there is no ambiguity, we write only $x \in G$ to denote that $x$ is in the vertex set of $G$. For $x, y \in G$ we denote by $x \sim_G y$ the neighbourhood relation of $G$. When no confusion arises, the subscript will be dropped and only the notation $\sim$ will be used.

All our graphs will have a special root vertex denoted by $o \in G$. For $x, y \in G$, let $\mathsf{d}(x, y)$ be the *graph metric*, that is, the length of the shortest path from $x$ to $y$. Also, write $d(x)$ for the degree of $x$ in $G$, i.e., the number of neighbours of $x$.

An *automorphism* of $G$ is a self-isometry of $G$ with respect to the graph metric $\mathsf{d}$, that is, a function $\varphi : V \to V$ which satisfies $\mathsf{d}(x, y) = \mathsf{d}\big(\varphi(x), \varphi(y)\big)$, for all $x, y \in V$.

We will need some fundamental notions from basic random walk theory. The presentation will mostly follow the notation from [Woe00].

Let $P = \big\{p(x, y)\big\}_{x,y\in G}$ be the transition probabilities of a random walk on $G$, which are adapted to the graph structure, i.e., $p(x, y) > 0$ if and only if $x$ is a neighbour of $y$. For $t \in \mathbb{N}$ we will denote by $p^t(x, y)$ the $t$ step transition probability, that is, the probability that the random walker goes from $x$ to

$y$ in exactly $t$ steps. A random walk is called *irreducible*, if for all vertices $x, y \in G$, there exists a $t$ such that $p^t(x, y) > 0$. We say a random walk is *reversible*, if there exists a measure $m : G \to \mathbb{R}$, such that

$$m(x)p(x, y) = m(y)p(y, x) \quad \text{for all } x, y \in G.$$

A measure $m$ with this property is called the *reversible measure* of the random walk $P$. Note that if $P$ is the *simple random walk* on $G$, that is

$$p(x, y) = \begin{cases} \frac{1}{d(x)}, & \text{for all } x \sim y \\ 0, & \text{otherwise,} \end{cases}$$

then $m(x) = d(x)$. All random walks under consideration will be irreducible and also reversible.

We will write $X_t$ for the position of the random walker at time $t$. Probabilities will be denoted by $\mathbb{P}$. If we write $\mathbb{P}_x$ for some $x \in G$, this will mean the probability of a random walk, which starts at vertex $x$. For random walks, which start at the root vertex $o$, we will often omit the subscript. Similarly $\mathbb{E}$ and $\mathbb{E}_x$ will denote expectations using the same convention.

**Definition 2.1.1** (Green function).

- *For $y, z \in G$ the* Green function *is defined as*

  $$G(y, z) = \mathbb{E}_y\Big[\sum_{t=0}^{\infty} \mathbf{1}_{\{X_t = z\}}\Big].$$

  *$G(y, z)$ is the expected number of visits to $z$ of the random walk $X_t$ started at $y$.*

- *For a subset $A \subset G$, write $G_A$ for the* Green function of the random walk stopped upon leaving the set $A$

  $$G_A(y, z) = \mathbb{E}_y\Big[\sum_{t=0}^{\tau-1} \mathbf{1}_{\{X_t = z\}}\Big],$$

  *with $\tau = \min\{t \geq 0 : X_t \notin A\}$.*

For a function $f : G \to \mathbb{R}$, its *Laplace operator* $\triangle f$ is defined as

$$\triangle f(x) = \sum_{y \sim x} p(x, y) f(y) - f(x).$$

**Definition 2.1.2.** *A function $f : G \to \mathbb{R}$ is called* superharmonic *on a set $B \subset G$, if $\triangle f(z) \leq 0$ for all $z \in B$ and* harmonic*, if $\triangle f(z) = 0$.*

**Lemma 2.1.3** (Minimum principle). *If $f$ is a superharmonic function on $G$ and there exists $x \in G$ such that $f(x) = \min_G f$, then $f$ is constant.*

## 2.2 Internal Aggregation Models

### 2.2.1 The Divisible Sandpile Model

The *divisible sandpile* model has been introduced by Levine and Peres [LP09] as a tool for studying internal growth models on $\mathbb{Z}^d$. It is a continuous analogue of the abelian sandpile model introduced by Bak, Tang and Wiesenfeld [BTW88].

Let $G$ be a graph, and $\mu_0$ a *mass distribution* on $G$, i.e., a function $\mu_0 : G \to \mathbb{R}_+$ with finite support, that is

$$\operatorname{supp}\mu_0 = |\{x \in G : \mu_0(x) > 0\}| < \infty.$$

The divisible sandpile is a sequence $\{\mu_k\}_{k\geq 0}$ of such mass distributions, which are created according to the following rule. At each time step $k$, choose a vertex $x$. If $\mu_k(x) \geq 1$ the sand at $x$ is unstable and topples, which means that the vertex $x$ keeps a mass of 1 for itself and the remaining mass $\mu_k(x) - 1$ is distributed among the neighbours of $x$ according to a given transition operator $p(x, y)$. If every vertex topples infinitely often, the mass distribution converges to a limit

$$\lim_{k\to\infty} \mu_k = \mu \leq 1.$$

The existence of this limit will be shown in Lemma 3.3.2 and in Lemma 3.3.3 for arbitrary locally finite graphs $G$, which may also have loops, and reversible transition operators.

Write $S$ for the set of fully occupied sites of the divisible sandpile with start distribution $\mu_0$, that is

$$S = \{x \in G : \ \mu(x) = 1\}. \tag{2.1}$$

We will show in Lemma 3.4.2 that $S$ does not depend on the order of the topplings as long as every vertex topples infinitely often. The set $S$ is called the *divisible sandpile cluster* for initial distribution $\mu_0$. Lemma 3.4.2 also gives a method for describing the set $S$ in terms of a discrete obstacle problem. We will use it to prove a shape theorem for the divisible sandpile on the *comb*.

The *comb*, which we will denote by $\mathcal{C}_2$, is the spanning tree of the two dimensional lattice $\mathbb{Z}^2$, which is obtained by removing all horizontal edges of $\mathbb{Z}^2$ except the ones on the $x$-axis.

**Definition 2.2.1.** *The comb $\mathcal{C}_2$ is the graph with vertex set $\mathbb{Z}^2$ and neighbourhood relation given by*

$$(x_1, y_1) \sim (x_2, y_2) \iff \begin{cases} x_1 = x_2 \wedge |y_1 - y_2| = 1, \\ |x_1 - x_2| = 1 \wedge y_1 = 0 \wedge y_2 = 0. \end{cases}$$

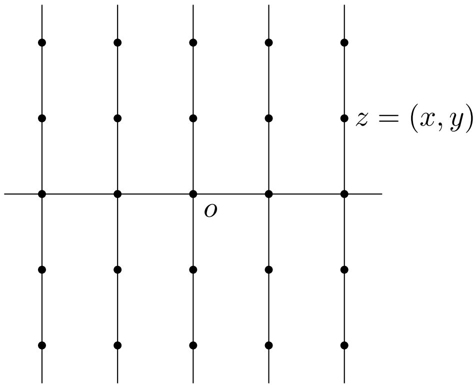



Figure 2.1: The two dimensional comb $\mathcal{C}_2$.

In other words, the graph $\mathcal{C}_2$ can be constructed from a two-sided inifinite path $\mathbb{Z}$ (the "*backbone*" of the comb), by attaching copies of $\mathbb{Z}$ (the "*teeth*") at every vertex of the backbone). We use the standard embedding of the comb into the two dimensional Euclidean lattice $\mathbb{Z}^2$, and use Cartesian coordinates $z = (x, y) \in \mathcal{C}_2$ to denote vertices of $\mathcal{C}_2$. The vertex $o = (0, 0)$ will be the root vertex, see Figure 2.1. For functions $g$ on the vertex set of $\mathcal{C}_2$ we will often for convenience write $g(x, y)$ instead of $g(z)$, when $z = (x, y)$.

Random walks on $\mathcal{C}_2$ have been studied by various authors, the first being Havlin and Weiss [WH86] and Gerl [Ger86]. While $\mathcal{C}_2$ is a very simple graph, it has some remarkable properties. For example, on the comb the so-called *Einstein relation* between the spectral-, walk- and fractal-dimension is violated, see Bertacchi [Ber06]. Peres and Krishnapur [KP04] showed that on $\mathcal{C}_2$ two independent simple random walks meet only finitely often; this is called the *finite collision property*. Further references include Bertacchi and Zucca [BZ03] and Csáki, Csörgő, Földes and Révész [CCFR09].

Let now $G = \mathcal{C}_2$, and $p(\cdot, \cdot)$ be the transition probabilities of the simple random walk on $\mathcal{C}_2$. As initial mass distribution we choose $\mu_0 = n \cdot \delta_o$, i.e. we start with mass $n$ concentrated at the origin and no mass anywhere else. Denote by $S_n$ the divisible sandpile cluster for this initial distribution. Using this setting we will prove the following shape theorem.

**Theorem 2.2.2.** *Let $S_n$ be the set of fully occupied sites for the divisible sandpile model on $\mathcal{C}_2$, with $\mu_0 = n \cdot \delta_o$, where $o = (0, 0)$ is the origin of the comb. Then there exists a constant $c \geq 0$, such that, for all $n \geq n_0$,*

$$\mathcal{B}_{n-c} \subset S_n \subset \mathcal{B}_{n+c},$$

*where*

$$\mathcal{B}_n = \left\{ (x,y) \in \mathcal{C}_2 : \frac{|x|}{k} + \left(\frac{|y|}{l}\right)^{1/2} \leq n^{1/3} \right\},$$

*and the constants $k$ and $l$ are given by $k = \left(\frac{3}{2}\right)^{2/3}$ and $l = \frac{1}{2}\left(\frac{3}{2}\right)^{1/3}$.*

### 2.2.2 Rotor-Router Aggregation

Rotor-router walks are deterministic analogues to random walks, which have been introduced into the physics literature under the name *Eulerian walks* by Priezzhev, D.Dhar et al [PDDK96] as a model of *self organized criticality*, a concept introduced by Bak, Tang and Wiesenfeld [BTW88].

In a rotor-router walk on a graph $G$, for each vertex $x \in G$ a cyclic ordering $c(x)$ of its neigbours is chosen, i.e., $c(x) = \big(x_0, \dots, x_{d(x)-1}\big)$, where $x_i \sim x$ and $x_i \neq x_j$. At each vertex we have an arrow (rotor) pointing to one of the neighbors of the vertex. When a particle is at a vertex $x$ two things happen. First the rotor is rotated to the next neighbour, defined by the ordering $c(x)$, and then the particle moves to that neighbour.

The behaviour of rotor-router walks is in some respects remarkably close to that of random walks. Cooper and Spencer [CS06] showed that rotor-router walks simulate random walks on $\mathbb{Z}^d$ with constant error, in the sense that they put a number of rotor-router particles at (apart from a technicality) arbitrary vertices and let the system evolve for a certain number of rounds. Each round every particle makes one step in a rotor-router walk (in arbitrary order), all sharing the same rotor configuration, and thus influencing the movement of each other. Cooper and Spencer showed that the difference between the number of rotor-router particles which are at vertex $v$ after $t$ rounds and the expected number of random walk particles which started in the same configuration after $t$ random walk steps is bounded by a constant $c$, which is independent of the number of particles, the number of rounds $t$ and also the vertex $v$. Doerr and Friedrich [DF06] improved this result in the case of $\mathbb{Z}^2$ and gave tight estimates for the constant $c$ depending on the selected rotor sequence.

Dimitriu, Tetali and Winkler [DTW03, Theorem 9.2] showed that on any finite tree $T$ the expected hitting time of random walk from vertex $u \in T$ to vertex $v \in T$

$$\mathbb{E}_u[\tau_v], \quad \text{with } \tau_v = \inf\{t \geq 0 : X_t = v\},$$

can be computed by a variant of the rotor-router model, in which the neighbour of a vertex $x$, which is the closest to the target vertex $v$, is always chosen last in the rotor sequence of $x$. In Chapter 5 we give a result in a similar spirit, which allows us to compute the *harmonic measure* of a finite subset $B$ of a graph $G$, that is, the probability

$$h(y) = \mathbb{P}_o[\tau_B = y], \quad \text{with } \tau_B = \inf\{t \geq 0 : X_t \notin B\},$$

in terms of a rotor-router process. This result is inspired by the methods of Holroyd and Propp [HP10], where they introduced a class of weight functions on the full configurations of the rotor-router model, that is, the position of all particles together with the state of all rotors. The weight functions are defined in such a way that the process of moving one or several particles, according to the rules of the rotor-router model, does not change the total weight of the system. Using this tool, they show that several quantities associated to the rotor-router walk, like hitting times and hitting probabilities, occupation times, etc., concentrate around their expected values for random walks. This approximation is also faster than in the case of random walks. For rotor-router walks the discrepancy after $n$ runs is of order $\mathcal{O}(n^{-1})$, while for random walks it is typically only of order $\mathcal{O}(n^{-1/2})$.

In the present work, we are mostly interested in **rotor-router aggregation**, which is analogous to internal DLA in the sense that the particles perform rotor-router walks instead of random walks. Let $R_1 = \{o\}$ and define the sets $R_n$ recursively, by

$$R_{n+1} = R_n \cup \{z_n\} \quad \text{for } n \geq 1,$$

where $z_n$ is the first vertex outside of $R_n$ that is visited by a rotor-router walk, started at $o$. Note that all rotor-router particles share the same rotor configuration, which is not reset when a new particle is started. We will call the set $R_n$ the *rotor-router cluster* of $n$ particles.

The *odometer function* $u(x)$ of the rotor-router aggregation is defined as the total number of times that some particle is sent out from vertex $x$ during the creation of the rotor-router cluster.

The first major contribution in the study of rotor-router aggregation was due to Levine and Peres [LP08]. They showed that on the lattice $\mathbb{Z}^d$ the set $R_n$, rescaled by $n^{-1/3}$, converges to the Euclidean unit ball in $\mathbb{R}^d$, in the sense that the Lebesgue measure of the symmetric difference of the rescaled rotor-router cluster and the unit ball goes to zero. In [LP09] the same authors improved on this result.

**Theorem 2.2.3** (Levine and Peres)**.** *Let $R_n$ be the rotor-router cluster on $\mathbb{Z}^d$ with $n$ particles starting at the origin, and let*

$$\mathcal{B}_r = \left\{x \in \mathbb{Z}^d : |x| < r\right\}$$

*be the Euclidean lattice ball of radius $r$. Then there exists constants $c_1, c_2$ depending only on $d$, such that*

$$\mathcal{B}_{r-c_1 \log r} \subset R_n \subset \mathcal{B}_{r(1+c_2 r^{-1/d} \log r)},$$

*where $r = \left(\frac{n}{\omega_d}\right)^{1/d}$, and $\omega_d$ is the volume of the unit ball in $\mathbb{R}^d$.*

While the bounds on the radii in Theorem 2.2.3 are much better than the best known estimates for IDLA on $\mathbb{Z}^d$, computer simulations as performed by Kleber [Kle05] and recently by Friedrich and Levine [FL] indicate that the errors are in fact much smaller, possibly even bounded independently of the number of particles $n$. Friedrich and Levine used an innovative algorithm which takes an approximation of the rotor-router odometer as its input. The runtime of the algorithm is then dependent on the accuracy of the initial guess. Using this technique they managed to compute rotor-router clusters with up to $2^{24}$ particles. For this cluster size the difference between the outer and inner radius is only 1.65.

Another interesting phenomenon is that the rotors arrange themselves into intricate patterns, whose general structure seems to be independent of the initial rotor configuration. On the other hand, the choice of the rotor sequence changes the appearence of the patterns drastically. See Figure 2.2 for two examples of rotor-router clusters on $\mathbb{Z}^2$ with different rotor sequences[1]. Still nothing is known about this patterns.

On homogeneous trees, rotor-router aggregation has been studied by Landau and Levine [LL09]. They showed that, if the initial rotor configuration is *acyclic*, that is, the directed subgraph of $G$ that is formed by the rotors containes no cycle, then the rotor-router cluster is a perfect ball with respect to the graph metric, whenever it has the right number of particles to form such a ball. For initial rotor configurations on that are not acyclic a shape theorem is not proved for homogeneous trees even though the method used in Chapter 4 of the present thesis should be applicable, but rather tedious.

Another example, where the rotor-router cluster is known explicitly, is the case of the *layered square lattice* $\hat{\mathbb{Z}}^2$. This is the multigraph obtained from $\mathbb{Z}^2$ by reflecting all edges on the coordinate axis away from the origin. For example, if $x > 0$, the vertex at position $(x, 0)$ has the four neighbours $(x, 1), (x, -1)$ and two times $(x + 1, 0)$. On this graph Kager and Levine [KL] showed that the resulting rotor-router cluster has, whenever possible, exactly the shape of a $L^1$ ball

$$D_r = \big\{(x, y) \in \hat{\mathbb{Z}}^2 : |x| + |y| \leq r\big\}.$$

[1] http://rotor-router.mpi-inf.mpg.de features high resolution images of rotor-router aggregates with up to 10 billion particles computed by Friedrich and Levine using the algorithm described in [FL].

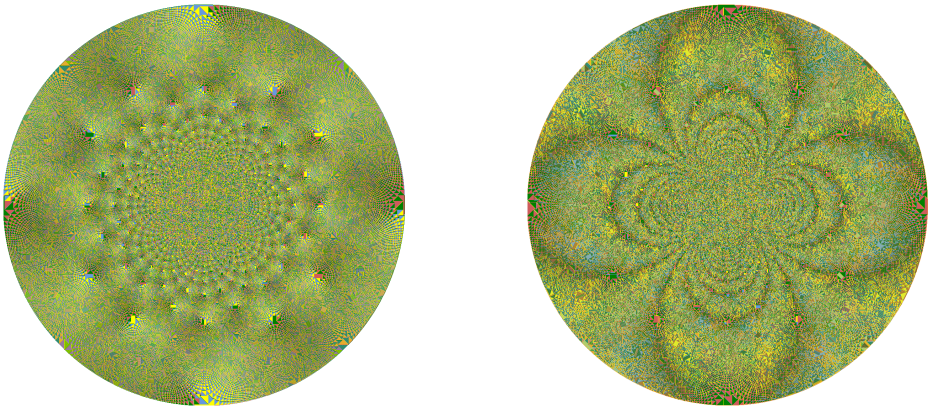


Figure 2.2: Two Rotor-router clusters mit $10^6$ particles on $\mathbb{Z}^2$. The four colors represent the different states of the rotors. The cluster on the left uses a cyclic rotor sequence $\{\uparrow, \rightarrow, \downarrow, \leftarrow\}$, while on the right an alternating sequence $\{\uparrow, \downarrow, \rightarrow, \leftarrow\}$ is used.

This result depends on a fixed rotor sequence and also on a fixed initial rotor configuration.

In this work we will focus on rotor-router aggregation on the comb. First we will use a similar method as in [KL] to show that, for a fixed initial rotor configuration, the rotor-router cluster, whenever it consists of the right number of particles, has exactly the shape

$$B_m = \big\{(x, y) \in \mathcal{C}_2 : \ |x| \leq m, |y| \leq h(m - |x|)\big\} \quad \text{for } m \in \mathbb{N},$$

with $h(x) = \left\lfloor \frac{(x+1)^2}{3} \right\rfloor$. This is the statement of Theorem 4.2.1.

For arbitrary initial rotor configurations and independently of the chosen rotor sequences, we can still say something, at least about an inner estimate for the aggregate. Using Theorem 2.2.2, and again with the help of rotor weights, we can show the following inner bound.

**Theorem 2.2.4.** *Let $R_n$ be the rotor-router cluster of $n$ particles on the comb. Then there exist constants $c_1, c_2$ and $c_3$, such that for $n \geq n_0$*

$$\tilde{\mathcal{B}}_n \subset R_n,$$

*where*

$$\tilde{\mathcal{B}}_n = \Big\{(x, y) \in \mathcal{C}_2 : \ |x| \leq kn^{1/3} - c_1 n^{1/6}, \\ |y| \leq l\left(n^{1/3} - \frac{x}{k}\right)^2 + c_2 x - c_3 n^{1/3}\Big\}.$$

### 2.2.3 Internal Diffusion Limited Aggregation

Let $\{X_t^n\}_{n\in\mathbb{N}}$ be a sequence of independent and identically $P$-distributed random walks on the graph $G$, with common starting point $X_0^n = o$.

**Definition 2.2.5.** *The* Internal Diffusion Limited Aggregation (IDLA) *is a stochastic process of increasing subsets $\{A_n\}_{n\in\mathbb{N}}$ of $G$, which are defined recursively as $A_1 = \{o\}$ and for $n \geq 2$*

$$\mathbb{P}\big[A_n = A_{n-1} \cup \{x\} | \, A_{n-1}\big] = \mathbb{P}_o\big[X_{\sigma_n}^n = x\big].$$

*Here $\sigma_n = \inf\{t \geq 0 : \ X_t^n \notin A_{n-1}\}$ is the stopping time of the first exit of the random walk $X_t^n$ from the set $A_{n-1}$.*

This means that at time $n$ a random walk $X_t^n$ is started at the root $o$, and evolves as long as it stays inside the IDLA-cluster $A_{n-1}$. When $X_t^n$ leaves $A_{n-1}$ for the first time, the random walk stops, and the point outside of the cluster that is visited by $X_t^n$ is added to the new cluster $A_n$.

This growth model was introduced by Diaconis and Fulton [DF91] in 1991. In 1992 Lawler, Bramson and Griffeath [LBG92] showed that for simple random walks on $\mathbb{Z}^d$, with $d \geq 2$, the limiting shape of IDLA (when properly rescaled) almost surely is the Euclidean ball of radius 1. In 1995 Lawler [Law95] refined this result by giving estimates on the fluctuations. More precisely, define the inner- and outer-fluctuation $\delta_I(n)$ and $\delta_O(n)$, such that the IDLA-cluster $A_{\lfloor \omega_d n^d \rfloor}$, where $\omega_d$ is the volume of the unit ball in $\mathbb{R}^d$, almost surely contains a ball of radius $n - \delta_I(n)$ and is contained in a ball of radius $n + \delta_O(n)$. Then the following estimates hold.

**Theorem 2.2.6** (Lawler, 1995)**.** *For IDLA on $\mathbb{Z}^d$ with $d \geq 2$, with probability* 1

$$\lim_{n\to\infty} \frac{\delta_I(n)}{n^{1/3}\log^2 n} = 0 \quad \textit{and} \quad \lim_{n\to\infty} \frac{\delta_O(n)}{n^{1/3}\log^4 n} = 0.$$

This bounds are not sharp. Computer simulations (see Moore and Machta [MM00] and Friedrich and Levine [FL]) suggest that the fluctuations are probably only logarithmic in the radius. Recently several improvements have been obtained. Asselah and Gaudillière [AG10a] and [AG10b] proved an upper bound of order $\log(n)$ for the inner fluctuation $\delta_I$ and of order $\log^2(n)$ for the outer fluctuation $\delta_O$ in all dimensions $d \geq 2$. Inpendently, and using different methods, Jerison, Levine and Sheffield [JLS10] proved that both $\delta_I$ and $\delta_O$ are of order $\log(n)$ for IDLA on $\mathbb{Z}^2$.

The case of IDLA on $\mathbb{Z}^d$ with simple random walk has been fully solved by Levine and Peres [LP10] (see also Levine [Lev07]), who showed that

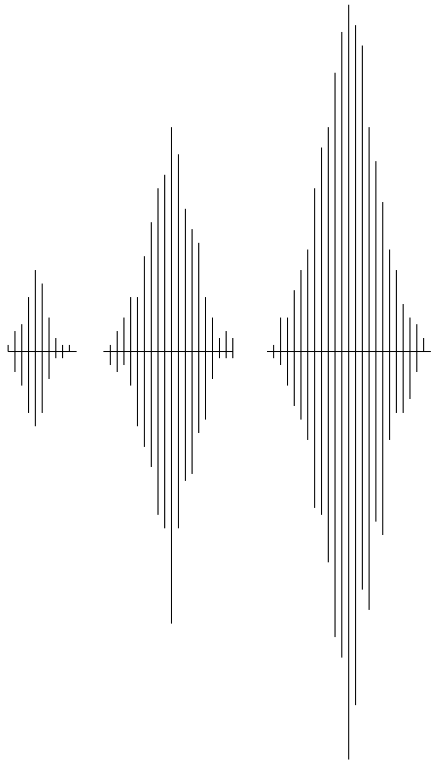


Figure 2.3: Internal DLA cluster on the comb for 100, 500 and 1000 particles.

IDLA, rotor-router aggregation as well as the divisible sandpile model all have the same scaling limit, when the lattice scaling tends to zero. This scaling limit can be described as the solution of a free boundary problem of a certain partial differential equation in $\mathbb{R}^d$.

Internal diffusion limited aggregation has also been studied in other contexts. Blachère [Bla00] proved a shape theorem for IDLA on $\mathbb{Z}^d$ for centered random walks, that is, random walks where the expectation of the increments $\mathbb{E}\big[X_t - X_{t-1}\big] = 0$, for all $t \geq 1$. Here the limiting shapes are always ellipsoids. In 2004 Blachère [Bla04] obtained a shape result for the IDLA model with symmetric random walk on discrete groups of polynomial growth, although with less precise bounds than in the case of $\mathbb{Z}^d$. Kager and Levine [KL10] proved that the IDLA clusters for a class of random walks on $\mathbb{Z}^2$, which they call *layered random walks*, have the limiting shape of $L^1$-balls.

Shellef [She10] investigated IDLA on the infinite cluster of supercritical Bernoulli bond percolation on $\mathbb{Z}^d$. Also in this case we have, for all $\varepsilon > 0$ and $n$ big enough,

$$\mathbb{P}\big[\mathcal{B}_{n(1-\varepsilon)} \subset A_{|\mathcal{B}_n|}\big] = 1.$$

Finally, in 2007 Blachère and Brofferio [BB07] proved a similar shape result for IDLA on finitely generated groups with exponential growth.

In all cases where IDLA dynamics has been studied so far, a common behaviour for the limiting shape of the clusters emerged. Namely, the limiting shape can be described as the level set of the Green function, that is, sets of the form

$$\{x \in G : \ G(o,x) \geq N\},$$

where $N$ is a positive constant.

This correspondence has been made particularly clear by Blachère and Brofferio [BB07], with the introduction of the *hitting distance* (also known as *Green metric*), a left invariant metric on finitely generated groups, which is defined as

$$d_H(x,y) = -\ln \mathbb{P}_x\big[\tau_y < \infty\big],$$

where $\tau_y = \inf\big\{t \geq 0 : \ X_t = y\big\}$ is the stopping time of the first visit to $y$. Further work on the hitting distance and its connections with random walk entropy and the Martin boundary has been done by Blachère, Haïssinsky and Mathieu [BHM08].

In this work we present two new results for IDLA. In Chapter 6 we consider IDLA for the simple random walk on the comb (see Figure 2.3), and we prove the following shape theorem.

**Theorem 2.2.7.** *Let $A_n$ be the internal DLA cluster of $n$ particles starting at the origin $o$, for the simple random walk on the comb $\mathcal{C}_2$. Then, for all $\varepsilon > 0$, we have with probability 1 that*

$$\mathcal{B}_{n(1-\varepsilon)} \subset A_n, \ \textit{for all sufficiently large } n \tag{2.2}$$

*where*

$$\mathcal{B}_n = \left\{ (x,y) \in \mathcal{C}_2 : \ \frac{|x|}{k} + \left(\frac{|y|}{l}\right)^{1/2} \leq n^{1/3} \right\}, \tag{2.3}$$

*and the constants $k$ and $l$ are given by $k = \left(\frac{3}{2}\right)^{2/3}$ and $l = \frac{1}{2}\left(\frac{3}{2}\right)^{1/3}$.*

Note that the set $\mathcal{B}_n$ here is the same as the limit set of the divisible sandpile model on the comb from Theorem 2.2.2.

The second new result is a shape theorem for IDLA on nonamenable graphs, where the underlying random walks are *strongly reversible* and *uniformly irreducible*, see Chapter 7 for details. In this case, IDLA behaves similarly to the case of groups with exponential growth, that was mentioned above. Also in this case the IDLA-clusters have the shape of level set of the Green function. For the precise statement see Theorem 7.2.1.

# Chapter 3

# The Divisible Sandpile Model

In this chapter we define the divisible sandpile model rigorously, and prove the existence of a limit which is independent of the order of topplings. Section 3.5 is devoted to the divisible sandpile on combs, in particular to the proof of Theorem 2.2.2.

## 3.1 The Odometer Function

Let us first define the divisible sandpile process in a rigorous way. For simplicity of notation, define the *reversed Laplacian* $\triangle'$ as

$$\triangle' f(x) = \sum_{y \sim x} p(y,x) f(y) - f(x) = m(x) \triangle f'(x),$$

where $f'(x) = \frac{f(x)}{m(x)}$ and $m(x)$ is the reversible measure of the transition operator $p(x,y)$.

Given a mass distribution $\mu_k$ and a vertex $x \in G$, the *toppling operator* is given as

$$T_x \mu_k = \mu_k + \alpha_k(x) \triangle' \delta_x,$$

where $\alpha_k(y) = \max\{\mu_k(y) - 1, 0\}$, for $y \in G$. Let now $\mu_0$ be the initial mass distribution, and $\{x_k\}_{k \geq 0}$ be a sequence of vertices in $G$ called the *toppling sequence*, which contains each vertex of $G$ infinitely often. Then we can define the mass distribution of the sandpile after $k$ steps as

$$\mu_{k+1}(y) = T_{x_k} \mu_k(y) = T_{x_k} \cdots T_{x_0} \mu_0(y).$$

This is the amount of mass present at $y$ after toppling the vertices $x_1, \ldots, x_k$ in succession. We distinguish several cases:

(a) For $y = x_k$:

$$\begin{aligned}
\mu_{k+1}(y) &= \mu_k(x_k) + \alpha_k(x_k) \, \triangle' \, \delta_{x_k}(x_k) \\
&= \mu_k(x_k) + \alpha_k(x_k) \left( \sum_{z \sim x_k} p(z, x_k)\delta_{x_k}(z) - \delta_{x_k}(x_k) \right) \\
&= \mu_k(x_k) + \max\{\mu_k(x_k) - 1, 0\}\big(p(x_k, x_k) - 1\big) \\
&= \begin{cases} 1 + p(x_k, x_k)\big(\mu_k(x_k) - 1\big), & \mu_k(x_k) \geq 1 \\ \mu_k(x_k), & \mu_k(x_k) < 1. \end{cases}
\end{aligned}$$

(b) For $y \sim x_k$:

$$\begin{aligned}
\mu_{k+1}(y) &= \mu_k(y) + \alpha_k(x_k) \, \triangle' \, \delta_{x_k}(y) \\
&= \mu_k(y) + \alpha_k(x_k) \left( \sum_{z \sim y} p(z, y)\delta_{x_k}(z) - \delta_{x_k}(y) \right) \\
&= \mu_k(y) + p(x_k, y) \max\{\mu_k(x_k) - 1, 0\} \\
&= \begin{cases} \mu_k(y) + p(x_k, y)\big(\mu_k(x_k) - 1\big), & \mu_k(x_k) \geq 1 \\ \mu_k(y), & \mu_k(x_k) < 1. \end{cases}
\end{aligned}$$

(c) For $y \neq x_k$ and $y \not\sim x_k$:

$$\mu_{k+1}(y) = \mu_k(y).$$

From (a) and (b) we get for $y \sim x_k$:

$$\begin{aligned}
\mu_{k+1}(x_k) - \mu_k(x_k) &= \big(p(x_k, x_k) - 1\big)\alpha_k(x_k) \\
\mu_{k+1}(y) - \mu_k(y) &= p(x_k, y)\alpha(x_k).
\end{aligned} \tag{3.1}$$

Therefore,

$$\sum_{y \in G} \big(\mu_{k+1}(y) - \mu_k(y)\big) = \big(p(x_k, x_k) - 1\big)\alpha_k(x_k) + \sum_{\substack{y \sim x_k \\ y \neq x_k}} p(x_k, y)\alpha_k(x_k) = 0,$$

and the sandpile does not leak mass. We can therefore define

$$M = \sum_{y \in G} \mu_0(y) = \sum_{y \in G} \mu_k(y) \tag{3.2}$$

as the *total amount of mass* of our sandpile.

The most important tool that will be used throughout this work in various incarnations is the so called *odometer function*, which was introduced by Levine and Peres [LP09].

**Definition 3.1.1.** *The* odometer function $v_k$ *is defined as*

$$v_k(y) = \sum_{j \leq k \colon x_j = y} \mu_j(y) - \mu_{j+1}(y) = \sum_{j \leq k \colon x_j = y} \alpha_j(y),$$

*and represents the total mass emitted from a vertex $y \in G$ during the first $k$ topplings.*

It turns out that it is easier to work with the odometer function than directly with the mass distributions, since it has some nice properties. It will be especially important in Chapters 4 and 6 where we use the results of this chapter in order to derive shape theorems for the rotor-router aggregation and IDLA.

In the following, we will state some results which hold for general reversible transition operators on arbitrary graphs $G$. For the case of $\mathbb{Z}^d$, the proofs can be found in LEVINE AND PERES [LP09]. The proof of the general case is included for completeness.

## 3.2 Convergence of the Odometer

In order to prove that the sequence of mass distributions $\mu_k$ has a limit, we first prove that the sequence of odometer functions $v_k$ converges.

**Lemma 3.2.1.** *As $k \to \infty$, the odometer function $v_k$ converges pointwise to a limit function $v$.*

*Proof.* Let

$$B = \big\{ y \in G : \mathsf{d}(y, \operatorname{supp} \mu_0) \leq M \big\}, \tag{3.3}$$

where $M$ is the total amount of mass of the sandpile, as defined in (3.2). We have $\operatorname{supp} \mu_k \subseteq B$ for all $k \geq 0$. Let now $\ell : G \to \mathbb{R}$ be such that $\triangle \ell(x) = 1$ for all $x \in B$. Define the weight

$$Q_k = \sum_{y \in G} \mu_k(y) \ell(y) \leq \max_{y \in B} \ell(y) \sum_{y \in G} \mu_k(y) = M \cdot \max_{y \in B} \ell(y)$$

From equation (3.1) we get

$$\begin{aligned} Q_{k+1} - Q_k &= \sum_{y \in G} \left( \mu_{k+1}(y) - \mu_k(y) \right) \ell(y) \\ &= (p(x_k, x_k) - 1) \alpha_k(x_k) \ell(x_k) + \sum_{\substack{y \sim x_k \\ y \neq x_k}} p(x_k, y) \alpha_k(x_k) \ell(y) \\ &= \alpha_k(x_k) \triangle \ell(x_k) = \alpha_k(x_k), \end{aligned}$$

where the last line uses the fact that $\alpha_k(x) = 0$ for $x \notin B$. Summing over $k$ gives

$$Q_{k+1} = Q_0 + \sum_{j=0}^{k} \alpha_j(x_j) = Q_0 + \sum_{y \in G} v_k(y) \leq \max_{y \in B} \ell(x) \cdot M.$$

Hence, for each $x \in G$ the odometer function $v_k(x)$ is a bounded function. Moreover, it is also increasing in $k$ and this implies the pointwise convergence of $v_k$. Denote by $v(x) = \lim_{k \to \infty} v_k(x)$ its pointwise limit. □

Note that in the previous proof we restrict the condition $\triangle \ell(x) = 1$ to the trivial finite upper bound $B$, to ensure the existence of the function $\ell$. For concrete calculations it is preferable to work with functions $\ell$ which satisfy the condition on all vertices of $G$, if possible. For example the function $\ell(x) = \|x\|^2$ does this on $\mathbb{Z}^d$ with simple random walk. Another example would be Cayley graphs of finitely generated groups where the existence of an function $\ell$ with $\triangle \ell(x) \equiv 1$ on the whole graph is ensured by a theorem of Ceccherini-Silberstein and Coornaert [CCS09].

## 3.3 Convergence of the Sandpile

Using the fact that the odometer function $v_k$ of the divisible sandpile converges pointwise to a limit function $v$, we can now prove the convergence of the mass distribution $\mu_k$ of the sandpile after $k$ topplings. The proof is again based on the result of Levine and Peres [LP09] for $\mathbb{Z}^d$. In the general case we need our assumption of reversibility of the transition operator $p(x, y)$. It will be convenient to work with a slightly modified odometer function, which takes the reversible measure into account.

**Definition 3.3.1.** *The* normalized odometer function $u : G \to \mathbb{R}_+$ *is defined as*

$$u(x) = \frac{v(x)}{m(x)},$$

*where $m(x)$ is the reversible measure of the underlying random walk on $G$ with transition matrix $P$.*

From now on we will mostly work with the normalized odometer function $u$. We are now able to prove the convergence of the mass distributions.

**Lemma 3.3.2.** *The mass distribution $\mu_k$ of the sandpile converges pointwise to a limit distribution $\mu$. Moreover*

$$\mu(x) = \mu_0(x) + m(x) \triangle u(x).$$

*Proof.* During the first $k$ topplings, a vertex $y \in G$ emits $p(y,x)v_k(y)$ of mass to a neighbour $x \sim y$. Therefore, the total amount of mass that $x$ receives during the first $k$ steps is equal to

$$\sum_{y \sim x} p(y,x)v_k(y).$$

Since the amount of mass present at $x$ at time $k$ is equal to the amount of mass at $x$ at the beginning, plus the amount of mass received by $x$ and minus the amount emmited by $x$, we have

$$\begin{aligned}
\mu_k(x) &= \mu_0(x) + \sum_{y \sim x} p(y,x)v_k(y) - v_k(x) \\
&= \mu_0(x) + m(x)\left(\sum_{y \sim x} p(x,y)\frac{v_k(y)}{m(y)} - \frac{v_k(x)}{m(x)}\right) \\
&= \mu_0(x) + m(x) \bigtriangleup u_k(x),
\end{aligned}$$

where $u_k(x) = \frac{v_k(x)}{m(x)}$. Using the convergence of $v_k$ proved in Lemma 3.2.1, we conclude that the mass distribution $\mu_k$ also converges, and the limit distribution $\mu$ satisfies the relation

$$\mu(x) = \lim_{k \to \infty} \mu_k(x) = \mu_0(x) + m(x) \bigtriangleup u(x).$$

□

If $G$ does not have loops, it is trivial that $\mu$ is bounded above by the constant function 1, since for every $y \in G$ we have $\mu_k \leq 1$ whenever $y = x_k$. Hence for each vertex $y$ there is a subsequence $\mu_{k_j}(y) \leq 1$. If $G$ has loops we have to be more careful.

**Lemma 3.3.3.** *The limit distribution $\mu$ of the sandpile satisfies*

$$\mu(x) \leq 1, \quad \textit{for all } x \in G.$$

*Proof.* Let us consider the set

$$C = \big\{x \in G : \exists \varepsilon_x > 0 \text{ s.t. } \mu_k(x) > 1 + \varepsilon_x \text{ for infinitely many } k\big\}.$$

Since $\mu_k(x) = 0$ for $x \notin B$, with $B$ defined as in (3.3), $C$ is a subset of the set $B$, and $C$ is finite. Assume that $C$ is not empty, which means that there exists some $x \in G$ with $\mu(x) > 1$.

Define the inner boundary $\partial_I C$ of $C$ as

$$\partial_I C = \big\{x \in C : \exists y \sim x, \text{ with } y \notin C\big\},$$

and choose vertices $x \in \partial_I C$, and $y \sim x$ with $y \notin C$. Since $x \in C$, the value of the odometer function $\mu_k(x)$ exeeds $1 + \varepsilon_x$ infinitely often. This implies that $\alpha_k(x) > \varepsilon_x$ infinitely often.

Because the toppling sequence is choosen such that every vertex topples infinitly often, $\varepsilon_x \cdot p(x, y)$ amount of mass is added to each neighbour $y \sim x$ infinitely often. Eventually $\mu_k(y)$ will be bigger than 1. Moreover,

$$\mu_k(y) > 1 + \frac{\varepsilon_x}{2} p(x, y)$$

infinitely often. Hence $y \in C$, which is a contradiction to the finiteness of $C$. So $\mu \leq 1$ on $G$. □

## 3.4 Abelian Property

Everything we did in the last three sections depends on the chosen toppling sequence $\{x_k\}_{k \geq 0}$. In the next Lemma we give a representation of the normalized odometer function $u$ which is independent of the toppling sequence. Since the limiting mass distribution depends only on the initial mass distribution and on the odometer function (Lemma 3.3.2), the limit behaviour of the divisible sandpile is independent of the toppling sequence. This is the *Abelian property*, which will be fundamental to our further investigations.

**Definition 3.4.1.** *Let $g : G \to \mathbb{R}$ be a function on $G$. Define its* least superharmonic majorant *on a set $\mathfrak{B} \subset G$ as:*

$$s_g^{\mathfrak{B}}(z) = \inf \big\{ f(z) : \ f \text{ superharmonic on } \mathfrak{B},\ f \geq g \big\}.$$

Remark that the function $s_g^{\mathfrak{B}}$ is itself superharmonic on $\mathfrak{B}$. From Lemma 3.3.2 and 3.3.3 we get

$$\triangle u(z) = \frac{1}{m(z)} \big( \mu(z) - \mu_0(z) \big) \leq \frac{1}{m(z)} \big( 1 - \mu_0(z) \big).$$

In particular, if $z \in S$, where $S$ is the set of fully occupied sites defined in (2.1), we have

$$\triangle u(z) = \frac{1}{m(z)} \big( 1 - \mu_0(z) \big). \tag{3.4}$$

Additionally, by definition, $u(z) = 0$, if $z \notin S$.

Let us consider a function $\gamma : G \to \mathbb{R}$, such that

$$\triangle \gamma(z) = \frac{1}{m(z)} \big( 1 - \mu_0(z) \big), \quad \text{for all } z \in \mathfrak{B}, \tag{3.5}$$

where $\mathfrak{B}$ is a subset of $G$ which contains the set of fully occupied sites $S$ of the sandpile, i.e. $\mathfrak{B} \supset B$ with $B$ defined as in (3.3).

**Lemma 3.4.2.** *The normalized odometer function $u$ of the sandpile satisfies*

$$u = (\gamma + s)\mathbf{1}_{\mathfrak{B}},$$

*where $s = s_{-\gamma}^{\mathfrak{B}}$ is the least superharmonic majorant of $-\gamma$, and*

$$\mathbf{1}_{\mathfrak{B}}(z) = \begin{cases} 1, & \text{if } z \in \mathfrak{B} \\ 0, & \text{otherwise}, \end{cases}$$

*is the indicator function of the set $\mathfrak{B}$.*

*Proof.* By the definition of the function $\gamma$, in equation (3.5), we know that $\triangle(u - \gamma)(z) \leq 0$ for $z \in \mathfrak{B}$. Therefore, $u - \gamma$ is superharmonic on $\mathfrak{B}$. Also $u$ is nonnegative on $\mathfrak{B}$ and this implies that $u - \gamma \geq -\gamma$ on $\mathfrak{B}$. Therefore $u - \gamma$ is a superharmonic majorant of $-\gamma$, which implies that $u \geq \gamma + s$ on the set $\mathfrak{B}$.

In order to prove that $u - \gamma \leq s$, let us consider the function $s + \gamma - u$, which is superharmonic on $S = \{z \in G : \mu(z) = 1\}$, because, for all $z \in S$, one has

$$\triangle(s + \gamma - u)(z) = \triangle s(z) \leq 0.$$

Outside the sandpile cluster $S$, $u(z) = 0$, and because $s$ is a majorant of $-\gamma$, we have $s+\gamma-u \geq 0$. By the minimum principle for superharmonic functions this inequality extends to the inside of $S$, hence $u \leq \gamma+s$. Therefore, $u = \gamma+s$ on $\mathfrak{B} \supset S$. □

**Remark 3.4.3** (Abelian property)**.** *The limit of the odometer function $u$ is independent of the toppling sequence $\{x_k\}_{k\geq 0}$, since Lemma 3.4.2 does not depend on any toppling sequence.*

**Remark 3.4.4.** *A consequence of the Abelian property is that $u$, $\mu$ and $S$ are invariant under all automorphisms of the graph $G$ which fix the start distribution $\mu_0$.*

With the help of Lemma 3.4.2, we shall next prove that the sandpile cluster on the comb has the shape described in Theorem 2.2.2.

## 3.5 Divisible Sandpile on Combs

We will now study the behavior of the divisible sandpile on $\mathcal{C}_2$. We only consider the case where the initial mass distribution is a point mass at the origin.

**Remark 3.5.1.** *Let $u_n$ be the normalized odometer function on $\mathcal{C}_2$ with respect to simple random walk, for the initial mass distribution $\mu_0 = n \cdot \delta_o$, and let $S_n$ be the sandpile cluster in this case. Then by* (3.4) *the Laplacian of the odometer function is given by*

$$\triangle u_n(z) = \frac{1}{d(z)}\big(1 - n \cdot \delta_o(z)\big), \tag{3.6}$$

*where $d(z)$ is the degree of $z$ in $\mathcal{C}_2$.*

We can reduce the odometer function $u_n$ on the comb to the odometer function of the divisible sandpile on $\mathbb{Z}$, which is easy to compute.

Let $\tilde{u}_n$ be the normalized odometer function of the divisible sandpile on $\mathbb{Z}$, with initial mass distribution concentrated at 0, that is, $\tilde{\mu}_0 = n \cdot \delta_0$. By Remark 3.4.3 it is clear that the sandpile cluster $\tilde{S}_n$ on $\mathbb{Z}$ in this case is a symmetric interval around the origin 0. To be precise, we have

$$\tilde{S}_n = \left[-\frac{n-1}{2}, \frac{n-1}{2}\right] \cap \mathbb{Z}.$$

To compute the odometer function $\tilde{u}_n$, by Lemma 3.4.2 we need to construct a function $\tilde{\gamma}_n : \mathbb{Z} \to \mathbb{R}$, such that

$$\triangle \tilde{\gamma}_n(y) = \frac{1}{2}\big(1 - n\delta_0(y)\big),$$

for all $y \in \mathbb{Z}$. One possible choice for $\tilde{\gamma}_n$ is

$$\tilde{\gamma}_n(y) = \frac{1}{2}\left(y - \frac{n}{2}\right)^2. \tag{3.7}$$

Since $\tilde{\gamma}_n$ is non-negative, the constant function 0 is a superharmonic majorant of $-\tilde{\gamma}_n$. Hence, by Lemma 3.4.2

$$\tilde{u}_n \le \tilde{\gamma}_n,$$

for all $n \geq 0$.

We can now go back to the comb. Consider $\gamma_n : \mathcal{C}_2 \to \mathbb{R}$ with

$$\gamma_n(x, y) = \tilde{\gamma}_{n_x}(y), \tag{3.8}$$

where $n_x \in \mathbb{R}$, for all $x \in \mathbb{Z}$.

It is clear that this is the right way to define $\gamma_n$, since we can interpret each number $n_x$ as the total amount of mass that ends up in the copy of $\mathbb{Z}$, which is attached to the vertex $x$. Each copy of $\mathbb{Z}$ in the comb runs its own sandpile, mostly independent of the others.

It is easy to check that $\gamma_n$ satisfies property (3.4) for all $z = (x, y) \in \mathcal{C}_2$, if and only if

$$n_x = n \cdot \mathbf{1}_{\{x=0\}} + \tilde{\gamma}_{n_{x-1}}(0) - 2\tilde{\gamma}_{n_x}(0) + \tilde{\gamma}_{n_{x+1}}(0) \tag{3.9}$$

holds for all $x \in \mathbb{Z}$.

If we now plug (3.7) into (3.9), and use the fact that $n_x = n_{-x}$ by symmetry (see Remark 3.4.3), we get the following explicit nonlinear recursion for the numbers $n_x$

$$n_0 = n + \frac{1}{4}n_1^2 - \frac{1}{4}n_0^2, \tag{3.10}$$

$$n_x = \frac{1}{8}n_{x-1}^2 - \frac{1}{4}n_x^2 + \frac{1}{8}n_{x+1}^2, \text{ for } x > 0. \tag{3.11}$$

Equation (3.11) has an explicit solution as a quadratic polynomial of the form

$$n_x = \frac{2}{3}x^2 - t \cdot x + \frac{9t^2 + 4}{24}, \tag{3.12}$$

where $t$ is a real parameter. The function $n_x$ is positive for all $x$. If we use the initial condition (3.10) for $n_x$, we get also an explicit equation for the parameter $t$

$$n = \frac{3}{16}t^3 + \frac{3}{4}t^2 + \frac{5}{12}t + \frac{1}{3}.$$

This equation has one real root, which is given by

$$t = T(n) + \frac{28}{27}T(n)^{-1} - \frac{4}{3}, \tag{3.13}$$

with

$$T(n) = \left( \frac{8\sqrt{2187n^2 - 2916n + 629}}{81\sqrt{3}} + \frac{24n - 16}{9} \right)^{\frac{1}{3}}.$$

By a series expansion around $n = \infty$ we get

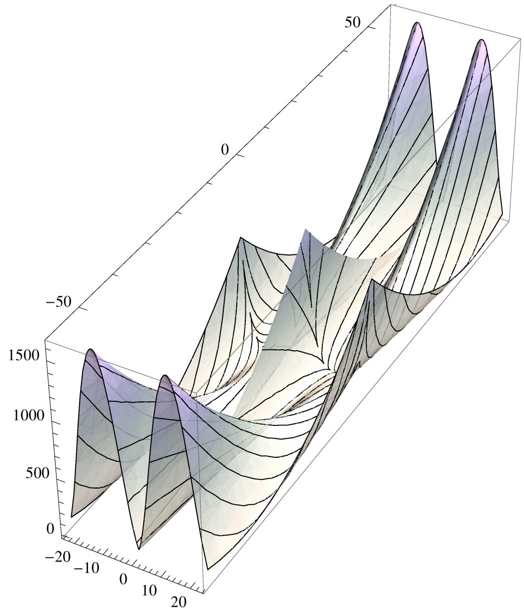
$$t = 2\left(\frac{2}{3}\right)^{1/3} n^{1/3} + \mathcal{O}(1). \tag{3.14}$$

Therefore, the function

$$\gamma_n(x, y) = \tilde{\gamma}_{n_x}(y) = \frac{1}{2}\left(y - \frac{n_x}{2}\right)^2, \tag{3.15}$$

with $n_x$ defined by (3.12) and (3.13) satisfies

$$\triangle\gamma_n(x, y) = \frac{1}{m(x)}\big(1 - \mu_0(x, y)\big)$$



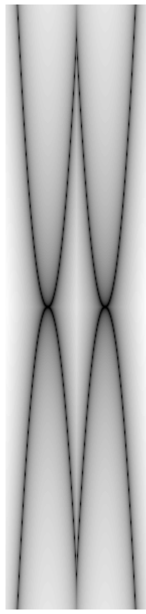



Figure 3.1: Two plots of the function $\gamma_n$ for $n = 1000$. The graphic on the left is superimposed with contour lines representing the the sets $\mathcal{B}_n$ for various values of $n$. In the density plot on the right, dark areas represent small values. By construction, the finite area which is surrounded by the local minima of $\gamma_n$ coincides with the region $S_n$ covered by the sandpile.

for all $(x, y) \in \mathcal{C}_2$.

With this in hand, we are now able to find an inner and outer estimate for $S_n$. This is the statement of Theorem 2.2.2. But first we need another simple fact about the odometer function $u_n$. For the proof see [LP09, Lemma 3.4].

**Lemma 3.5.2.** *If $x \in S_n \setminus \{o\}$ and $y \sim x$ with $\mathsf{d}(o, y) < \mathsf{d}(o, x)$, then*

$$u_n(y) \geq u_n(x) + 1.$$

***Proof of Theorem 2.2.2.***
*The upper bound:* The mass distributions $n_x$ are nonnegative for all $x$, therefore $\gamma_n$ is nonnegative, which implies that the constant function 0 is a superharmonic majorant of $-\gamma_n$. Hence, it follows by Lemma 3.4.2, that $\gamma_n$ is an upper bound of the odometer function $u_n$.

Since $u_n$ decreases by a fixed amount (Lemma 3.5.2) on the set $S_n$ of fully occupied sites when we move away from the origin, to get an upper bound of

$S_n$, it suffices to calculate the minima of $\gamma_n$ along each infinite ray starting at $o$. By Remark 3.4.4 it suffices to consider only the first quadrant.

First we consider the two rays which lie entirely on the $x$-axis, where $\gamma_n(x,0) = \frac{1}{8}n_x^2$. The minimum of this function is attained at position $x_{\min} = \frac{3}{4}t$, with $t$ given as in (3.13). Using the series expansion (3.14) of $t$ we get

$$x_{\min} = kn^{1/3} + \mathcal{O}(1), \qquad \text{with } k = \left(\frac{3}{2}\right)^{2/3}, \tag{3.16}$$

which is also an upper bound of $S_n$ on the $x$-axis by Lemma 3.5.2, since $\gamma_n(\lfloor x_{\min}\rfloor, 0)$ is bounded by a constant which is independent of $n$ and smaller than $1/10$.

To calculate the extent of the sandpile cluster on the "teeth", we need to compute the minima of $\gamma_n$ in the $y$-direction. On each "tooth" of the comb $\gamma_n$ is a quadratic polynomial which has its minimum at $y_{\min}(x) = \frac{n_x}{2}$. Moreover $\gamma\big(x, \lfloor y_{\min}(x)\rfloor\big) \leq \frac{1}{2}$. Using (3.12) and a series expansion around infinity we get

$$y_{\min}(x) = l\left(n^{1/3} - \frac{x}{k}\right)^2 + \frac{2}{3}x - \frac{1}{2l}n^{1/3} - \frac{7l}{9k}xn^{-1/3} + \mathcal{O}(1),$$

where $l = \frac{1}{2}\left(\frac{3}{2}\right)^{1/3}$. By the estimate in the $x$-direction we know that $(x,y) \in S_n$ only if $x \leq x_{\min}$, hence, using the expansion (3.16) for $x_{\min}$, we obtain

$$y_{\min}(x) \leq l\left(n^{1/3} - \frac{x}{k}\right)^2 + \mathcal{O}(1). \tag{3.17}$$

So for $n \geq n_0$, the following two inequalties hold for $(x,y) \in S_n$

$$\begin{aligned} |x| &\leq kn^{1/3} + \mathcal{O}(1), \\ |y| &\leq l\left(n^{1/3} - \frac{x}{k}\right)^2 + \mathcal{O}(1). \end{aligned}$$

This proves the outer estimate of Theorem 2.2.2, that is, $S_n \subset \mathcal{B}_{n+c}$.

*The lower bound:* In the previous part of the proof we have seen that on each infinite ray the minimum of $\gamma_n(z)$ is smaller than a non-negative constant $a$, which is independent of $n$.

By the outer estimate we know that $u_n(z) = 0$ for all $z \in \partial\mathcal{B}_{n+c}$. Hence $u_n(z) - \gamma_n(z) \geq -a$ for all $z \in \partial\mathcal{B}_{n+c}$. Since the function $u_n - \gamma_n$ is superharmonic, by the *Minimum Principle*, it attains its minimum on the boundary and the inequality

$$u_n(z) - \gamma_n(z) \geq -a$$

holds for all $z \in \mathcal{B}_{n+c}$. Thus $\gamma_n - a$ is also a lower bound of the odometer function on $\mathcal{B}_{n+c}$, which gives the inner estimate $\mathcal{B}_{n-c} \subset S_n$, for some constant $c$. □

In the next chapters we will use the odometer function $u_n$ in order to derive shape theorems for IDLA and the rotor-router model. For further reference we formulate the following corrollary.

**Corollary 3.5.3.** *Let $u_n$ be the normalized odometer function of the divisible sandpile on the comb $\mathcal{C}_2$, with initial mass distribution $\mu_0 = n \cdot \delta_o$, and $\mathcal{B}_n$ the subset of $\mathcal{C}_2$ as defined in Theorem 2.2.2. Then there exists a constant $0 < a < 2$, such that for all $n > n_0$ and all $z \in \mathcal{C}_2$*

$$\big(\gamma_n(z) - a\big)\mathbf{1}_{\mathcal{B}_n} \le u_n(z) \le \gamma_n(z).$$

Note that the same method could, at least in priciple, be applied to generalized combs which are defined as follows, see also [KP04].

**Definition 3.5.4.** *Given two graphs $G$ and $H$ with neighbourhood relations $\sim_G$ and $\sim_H$, and a root vertex $o \in H$, define $\mathbf{Comb_v}(G,H)$ to be the graph with vertex set $V(G) \times V(H)$ and neighbourhood relation*

$$(g_1,h_1) \sim (g_2,h_2) \Leftrightarrow \big[g_1 = g_2 \text{ and } h_1 \sim_H h_2\big] \text{ or } \big[g_1 \sim_G g_2 \text{ and } h_1 = h_2 = o\big].$$

If $G$ and $H$ are both recurrent, also $\mathbf{Comb_v}(G,H)$ is a recurrent graph. Peres and Krishnapur [KP04] showed that in this case the finite collision property also holds for these generalized combs.

Assuming we know the odometer function $u_n^H$ for the divisible sandpile on $H$ with initial mass distribution $\mu_0^H = n \cdot \delta_o$, one can analogously to (3.8) define the odometer on $\mathbf{Comb_v}(G,H)$ as

$$u(g,h) = u^H_{n_g}(h),$$

as long as we restrict ourselves to an initial distribution $\mu_0$ which is concentrated on $G$. Again $n_g$ is the amount of mass that ends up in the copy of $H$, which is attached to $G$ at vertex $g$. Hence $n_g \ge 0$, and they should satisfy the equation

$$\triangle^G \bar{u}(g) = \frac{1}{d_G(g)}\big(n_g - \mu_0(g,o)\big), \tag{3.18}$$

for all $g \in G$, for which the vertex $(g,0)$ is in the sandpile cluster. Here $\triangle^G$ is the Laplace operator on $G$ and the function $\bar{u} : G \to \mathbb{R}$ is defined as

$$\bar{u}(g) = u^H_{n_g}(o).$$

Equation (3.9) is a special case of (3.18) for the comb $\mathcal{C}_2 = \mathbf{Comb_0}(\mathbb{Z},\mathbb{Z})$. Unfortunately, already in the next easiest case, the iterated 3-dimensional comb $\mathcal{C}_3 := \mathbf{Comb_0}(\mathcal{C}_2,\mathbb{Z})$, equation (3.18) is quite hard to work with.

## Chapter 4

# Rotor Router Aggregation

In this chapter we prove two shape theorems for rotor-router aggregation on the comb $\mathcal{C}_2$. The first result, Theorem 4.2.1, gives an exact description of the rotor-router cluster for a specific initial rotor configuration (see Figure 4.1), and clockwise rotor sequence for all vertices. The second, Theorem 4.3.4, gives a weaker inner bound for arbitrary initial rotor configurations. Finally, in Section 4.4 we prove an inner bound for rotor-router aggregation, that is valid on all regular graphs.

## 4.1 The Abelian property

A *rotor configuration* on $G$ is a function

$$\rho : G \to G,$$

with $\rho(x) \sim x$, for all $x \in G$. Hence $\rho$ assigns to every vertex one of its neighbours. A rotor configuration $\rho$ is called *acyclic*, if the subgraph of $G$ spanned by the rotors contains no directed cycles. A particle configuration on $G$ is a function $\sigma : V \to \mathbb{Z}$, with finite support. If $\sigma(x) = m > 0$, we say that there are $m$ particles at vertex $x$.

The *rotor sequence* at vertex $x$ will be denoted by $c(x) = \big(x_0, x_1, \dots, x_{d(x)-1}\big)$ where all $x_i \sim x$ and $x_i \neq x_j$ for $i \neq j$.

**Definition 4.1.1** (Toppling operator)**.** *For a rotor configuration $\rho$ and a particle configuration $\sigma$, we define the* toppling operator $F_v$*, which sends one particle out of vertex $v$, by*

$$F_v(\rho, \sigma) = (\rho', \sigma'),$$

*where*

$$\rho'(w) = \begin{cases} \rho(w)^+ & \text{if } w = v, \\ \rho(v) & \text{otherwise.} \end{cases}$$

*Here* $\rho(w)^+$ *is defined as*

$$\rho(w)^+ = w_{(i+1) \bmod d(w)},$$

*if* $\rho(w)$ *equals* $w_i$ *in the cyclic ordering* $c(w)$ *of* $w$*, and the new particle configuration is given as*

$$\sigma'(w) = \begin{cases} \sigma(w) - 1 & \text{if } w = v, \\ \sigma(w) + 1 & \text{if } w = \rho'(v), \\ \sigma(w) & \text{otherwise.} \end{cases}$$

So $F_v$ first changes the rotor configuration by rotating the arrow at $v$ to its next position, and then it sends a particle along the edge the rotor at $v$ is now pointing at. The operation $F_v$ of toppling at some vertex $v$ can be successful even if there is no particle at $v$. If this is the case, then a "virtual particle" is sent away from $v$ and a "hole" left there. If there is already a hole at $v$, the operator $F_v$ will increase its depth by one. In the normal rotor-router aggregation no holes are allowed to be created during the whole process. A sequence of topplings $\{v_k\}_k \geq 1$ is called *legal*, if no holes are created when the vertices $v_k$ are toppled in sequence.

Note that the toppling operators commute, i.e., $F_v F_w = F_w F_v$ for all $v, w \in G$. This is the usual abelian property for rotor-router walks. While the final configuration result is always the same, rearranging the order of the topplings can turn a legal toppling sequence, into one that creates holes and virtual particles.

Given a function $u : G \to \mathbb{N}$ we denote

$$F^u = \prod_{v \in G} F_v^{u(v)},$$

where product means composition of the operators. Because of the abelian property, $F^u$ is well defined.

The proof of Theorem 4.2.1 is an application of a stronger version of the usual Abelian property of rotor-router walks, which has been recently introduced by Kager and Levine [KL]. We state it here for completness. For the proof see [KL].

**Theorem 4.1.2** (Strong Abelian Property)**.** *Let* $\rho$ *be a rotor configuration and* $\sigma$ *a particle configuration on* $G$*. Given two functions* $u_1, u_2 : G \to \mathbb{N}$*, write*

$$F^{u_i}(\rho, \sigma) = (\rho_i, \sigma_i), \qquad i = 1, 2.$$

*If* $\sigma_1 = \sigma_2$ *on* $G$*, and* $\rho_1$ *and* $\rho_2$ *are both acyclic, then* $u_1 = u_2$*.*

By the strong Abelian property we know that each end configuration can only be achieved by an unique amount of topplings for each vertex, even if we allow virtual particles to be formed during the process. In the case of rotor-router aggregation we have $\sigma = n \cdot \delta_o$. The desired end configuration is equal to $\sigma_1(x) = \mathbf{1}_{\{x \in R_n\}}$.

Thus, as soon as we find a valid odometer function $u_1(x)$ for the end configuration $\sigma_1(x)$, we also know that there exists a legal toppling sequence where each vertex $x$ topples exactly $u_1(x)$ times, and which does not create virtual particles during the process.

Friedrich and Levine [FL] use this principle to give a exact characterization of the odometer function of rotor-router aggregation.

**Theorem 4.1.3** (Friedrich, Levine)**.** *Let $G$ be a finite or infinite directed graph, $\rho_0$ a initial rotor configuration on $G$, and $\sigma_0$ a particle configuration on $G$. Let $u$ be the rotor-router odometer function of $\sigma_0$.*

*Fix $u_\star : G \to \mathbb{N}$, and let $A_\star = \{x \in G : u_\star(x) > 0\}$. Further define $\rho_\star$ and $\sigma_\star$ by*

$$F^{u_\star}(\rho_0, \sigma_0) = (\rho_\star, \sigma_\star).$$

*Suppose the following properties hold:*

- $\sigma_\star \leq 1$
- *$A_\star$ is finite*
- *$\sigma_\star(x) = 1$ for all $x \in A_\star$*
- *$\rho_\star$ is acyclic on $A_\star$.*

*Then $u_\star = u$.*

## 4.2 Specific Initial Rotor Configuration

The next result gives the exact shape of rotor-router aggregation on the comb, for a specific initial rotor configuration and fixed rotor sequence.

**Theorem 4.2.1.** *Let $R_n$ be the rotor-router cluster of $n$ particles on the comb $\mathcal{C}_2$, with initial rotor configuration $\rho_0$, defined as in Figure 4.1, and clockwise rotor sequence for all $x \in \mathcal{C}_2$. Define*

$$B_m = \{(x, y) \in \mathcal{C}_2 : \ |x| \leq m, |y| \leq h(m - |x|)\} \quad \textit{for } m \in \mathbb{N}, \tag{4.1}$$

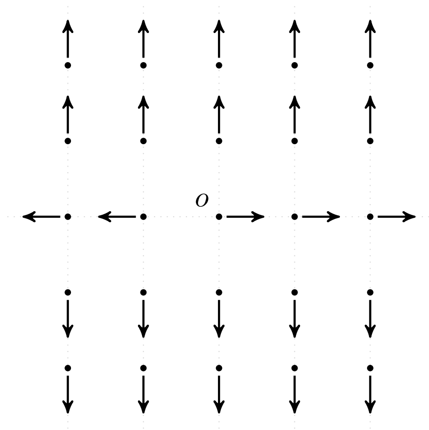


Figure 4.1: The initial acyclic rotor configuration $\rho_0$.

*with* $h(x) = \left\lfloor \frac{(x+1)^2}{3} \right\rfloor$. *Then for* $n_m = |B_m|$, *the rotor-router cluster satisfies*

$$R_{n_m} = B_m, \qquad \text{for all } m \geq 0.$$

Note that the set $B_m$ coincides with the set $\mathcal{B}_n$ from Theorem 2.2.2, with the exception of two points on the $x$-axis, whenever $m \in \mathbb{N}$ and $n \in \mathbb{R}_{\geq 0}$ are chosen such that $m = kn^{1/3} + 1$, with $k = \left(\frac{3}{2}\right)^{2/3}$. As a matter of fact, if we parameterize $\mathcal{B}_n$ by the $x$-coordinate of its rightmost point on the $x$-axis instead by its area, we get the following

$$\mathcal{B}_n = \tilde{\mathcal{B}}_{\tilde{m}} = \big\{(x, y) \in \mathcal{C}_2 : |x| \leq \tilde{m} \text{ and } |y| \leq \tilde{h}(\tilde{m} - |x|)\big\},$$

with $\tilde{m} = kn^{1/3}$ and $\tilde{h}(x) = \frac{x^2}{3}$. Hence, for all $m \in \mathbb{N}$,

$$\tilde{\mathcal{B}}_{m+1} = B_m \cup \big\{(-m-1, 0), (m+1, 0)\big\}.$$

For the proof of Theorem 4.2.1, an exact expression for the cardinality of the sets $B_m$ is needed.

**Proposition 4.2.2.** *Let* $B_m$ *be the set defined in equation* (4.1). *Then, for all* $m \geq 0$ *the cardinality of* $B_m$ *is given by*

$$|B_m| = \frac{1}{9}\big(4m^3 + 12m^2 + 24m + 5 + 2((m+2) \bmod 3)\big). \tag{4.2}$$

*Proof.* In order to simplify the statement of the proposition, we have to distinguish three cases, namely for $m = 3k + i$, with $i = 0, 1, 2$. The right-hand side of (4.2) is then equal to

$$\begin{aligned} N_0(k) &= 12k^3 + 12k^2 + 8k + 1, \text{ for } m = 3k \\ N_1(k) &= 12k^3 + 24k^2 + 20k + 5, \text{ for } m = 3k + 1 \\ N_2(k) &= 12k^3 + 36k^2 + 40k + 15, \text{ for } m = 3k + 2. \end{aligned} \tag{4.3}$$

We prove (4.3) by induction over $k$. The base case of the induction ($m = 0$) is immediate from the definition of $B_m$. The inductive step follows from

$$|B_m| = |B_{m-1}| + 2\big(h(m) + h(m+1) + 1\big).$$

□

We will use Theorem 4.1.3 in order to prove an exact formula for the odometer function of the rotor-router aggregation defined in Theorem 4.2.1. For this, we first look at the router-router process on the non-negative integers.

### 4.2.1 Rotor-Router on the non-negative Integers

For a better understanding of the rotor-router process on $\mathcal{C}_2$, we first analyse it on its "half-teeth", that is, on the nonnegative integers, where it is very simple. Consider $G = \mathbb{N}_0$, where the vertex 0 is a sink, and the initial rotor configuration $\rho_0$ is given by

$$\rho_0(y) = y + 1, \quad \forall y \geq 1.$$

Let $\tilde{R}_0 = \{1\}$, and define a modified rotor-router aggregation process $\tilde{R}_n$ recursively as follows. Start a rotor walk in 1, and stop the particle when it either reaches the sink 0, or exits the previous cluster $\tilde{R}_{n-1}$. Denote by $\tilde{z}_n$ the vertex where the $n$-th particle stops, and by $\rho_n$ and $\tilde{u}_n$ the rotor configuration and odometer function at that time. Then,

$$\tilde{R}_n = \begin{cases} \tilde{R}_{n-1} \cup \{\tilde{z}_n\}, & \text{if } \tilde{z}_n \neq 0 \\ \tilde{R}_{n-1}, & \text{otherwise.} \end{cases}$$

Obviously $\tilde{R}_n = \{1, \ldots, h_n\}$ for some sequence $h_n$. Since $\rho_0$ is acyclic, all rotor configurations $\rho_n$ are acyclic and have the form

$$\rho_n(y) = \begin{cases} y - 1, & 0 \leq y \leq r_n \\ y + 1, & \text{otherwise,} \end{cases}$$

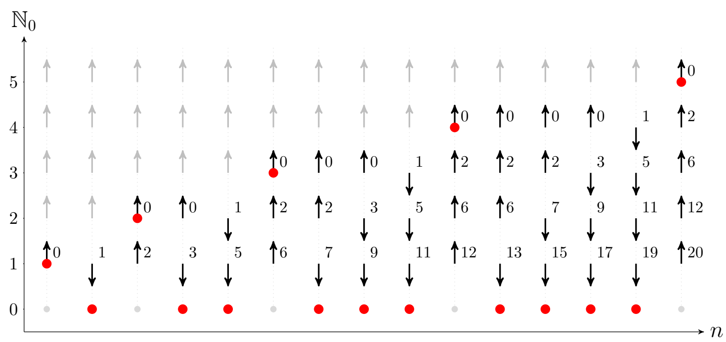


Figure 4.2: The first steps of the process $\tilde{R}_n$ on $\mathbb{N}_0$. The dots mark the vertex, where the current particle stopped.

for some numbers $0 \leq r_n \leq h_n$.

The odometer function $\tilde{u}_n$ is given by

$$\tilde{u}_n(y) = \tilde{u}(h_n, r_n, y) = \begin{cases} f(h_n - y) + e(r_n - y)\mathbf{1}_{\{y \leq r_n\}}, & 1 \leq y \leq h_n \\ 0, & \text{otherwise}, \end{cases} \tag{4.4}$$

where $e(y) = 2y + 1$ and $f(y) = y(y + 1)$. That $\tilde{u}_n$ correctly describes the odometer function of $\tilde{R}_n$ can be easily verified by induction. See Figure 4.2 for a graphical representation of the process $\tilde{R}_n$.

### 4.2.2 Odometer on the Comb

Since the rotor-router aggregation on the "teeth" of $\mathcal{C}_2$ behaves like the process $\tilde{R}_n$ from the previous section, it is enough to determine the numbers $r_n$ and $h_n$ in (4.4), depending on $x$, in order to fully specify the odometer function on $\mathcal{C}_2$ for points off the $x$-axis.

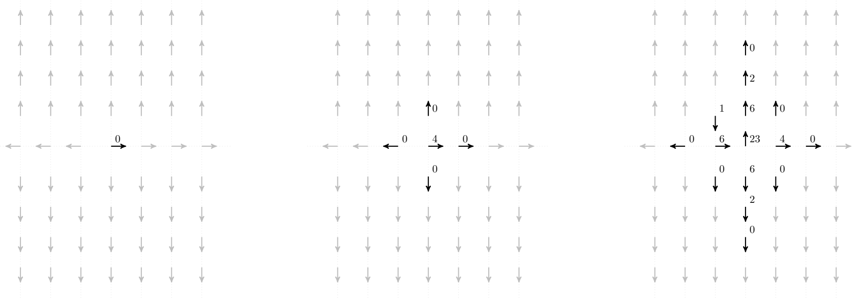
Consider the sequences given by

$$r_x = \begin{cases} 0, & x \in \{0, 1\} \\ \frac{1}{18}\left(x^2 - 7x + 10\right), & x \equiv 2 \bmod 3 \\ \frac{1}{6}\left(x^2 - x + 6\right), & \text{otherwise}, \end{cases} \tag{4.5}$$

and $h_x = \left\lfloor \frac{(x+1)^2}{3} \right\rfloor$ as in the definition of $B_m$.

Define $u_m : B_m \to \mathbb{N}$ by

$$u_m(x, y) = u'(m - |x|, |y|), \tag{4.6}$$



$n = 1$ $n = 5$ $n = 15$

Figure 4.3: The first three fully symmetric configurations, consisting of $n$ particles. The numbers are the values of the odometer function $u_n$.

where

$$u'(x, y) = \begin{cases} \tilde{u}\big(h_x, r_x, y\big), & y > 0 \\ 2\tilde{u}\big(h_x, r_x, y\big) - 2 - \mathbf{1}_{\{x=2\}}, & y = 0, \end{cases}$$

with $\tilde{u}$ as in (4.4). We claim that $u_m$ is the odometer function for rotor-router aggregation of $|B_m|$ particles on the comb.

### 4.2.3 Proof of Theorem 4.2.1

**Definition 4.2.3.** *Let $(\rho, \sigma)$ be the final configuration of the rotor-router aggregation process described in Theorem 4.2.1 for $|B_m|$ particles. The configuration $(\rho, \sigma)$ is then called the $m$-th* fully symmetric configuration.

Figures 4.3 and 4.4 show examples of fully symmetric configurations.

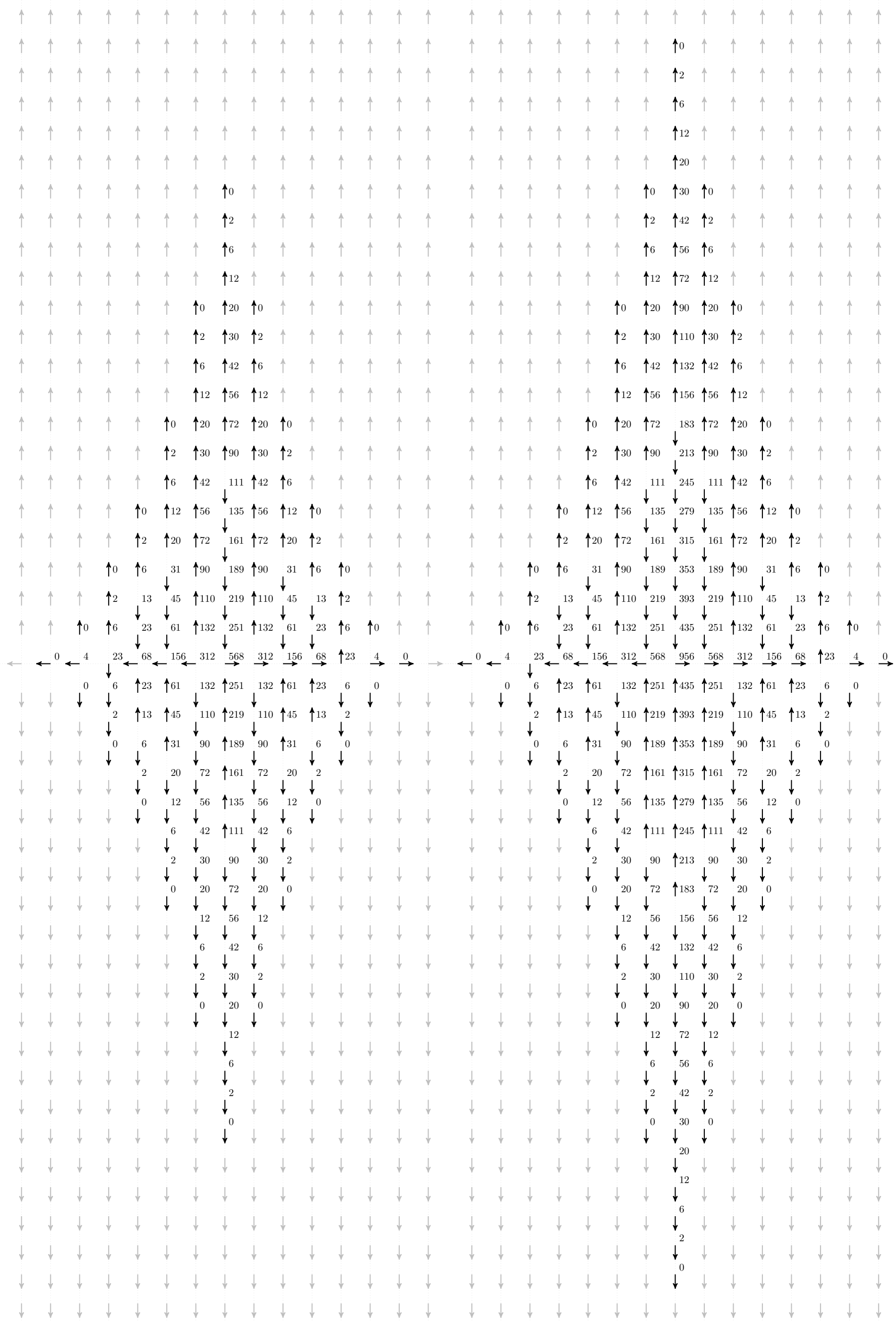
Next we will show that, if $m$ is big enough, the toppling function $u_m$ defined in (4.6) generates exactly the $m$-th fully symmetric configuration, if $|B_m|$ particles start at the origin.

**Lemma 4.2.4.** *Let $\rho_0$ be the initial rotor configuration defined in Figure 4.1, and $\sigma_0 = |B_m|\delta_o$. Furthermore, define $\rho_m$ and $\sigma_m$ as*

$$(\rho_m, \sigma_m) = F^{u_m}(\rho_0, \sigma_0),$$

*where $u_m$ is defined as in (4.6). A clockwise rotor sequence is assumed for all vertices. Then $\sigma_m = \mathbf{1}_{B_m}$, for all $m \geq 2$.*

*Proof.* We can use Theorem 4.1.3 to verify that $u_m$ is indeed the odometer function of this rotor-router process. For this we need to check the four



$n = 161$ $n = 237$

Figure 4.4: The $6^{\text{th}}$ and $7^{\text{th}}$ fully symmetric configurations, consisting of $n$ particles. The numbers are the values of the odometer function $u_n$.

properties of Theorem 4.1.3, with

$$A_\star = B_m \setminus \big\{z \in B_m : \exists y \sim z \text{ such that } y \notin B_m\big\}.$$

The set $A_\star$ is obviously finite. For those vertices $z \in A_\star$ that have neighbours in $B_m \setminus A_\star$, we have, by (4.4), $u_m(z) \le 3$, if $z$ is not on the $x$-axis and $u_m(z) = 6$ otherwise. In both cases exactly one particle is sent to the vertex outside of $A_\star$, hence $\sigma_m(z) \le 1$ for all $z \notin A_\star$.

Next we prove that $\sigma_m(z) = 1$, for all $z \in A_\star$. Since by definition $u_m$ is symmetric, we need to consider only one quadrant. Let $z = (x, y) \in B_m$ with $x, y \ge 0$. We need to distinguish several cases.

**Case 1.** $y \ge 2$: This is the case of the rotor-router aggregation on the nonnegative integers. The number of particles the vertex $z$ receives from its neighbours should be one more than the number of particles $z$ sends to its neighbours. From Figure 4.2, and due to the fact that the final rotor configuration restricted to each "tooth" is acyclic, there are only four possible situations:

(a) The rotors at the vertices $(x, y-1)$, $(x, y)$ and $(x, y+1)$ all point outwards ($\uparrow$). This is the case when $r_x < y - 1$, hence the vertex $(x, y)$ receives $\frac{1}{2}\tilde{u}_m(y-1) + \frac{1}{2}\tilde{u}_m(y+1)$ particles from its right and left neighbours, and it sends $\tilde{u}_m(y)$ particles. That is,

$$\sigma_m(x, y) = \frac{1}{2}\big[f(h_x - y + 1) + f(h_x - y - 1)\big] - f(h_x - y) = 1.$$

(b) The rotors at the vertices $(x, y-1)$, $(x, y)$ and $(x, y+1)$ all point inwards ($\downarrow$). Hence $r_x \ge y + 1$ and, comparing the numbers of incoming and outgoing particles, we have

$$\begin{aligned}\sigma_m(x, y) = \frac{1}{2}\big[&f(h_x - y + 1) + e\big(r_x - y + 1) + f(h_x - y - 1)\\ &+ e(r_x - y - 1)\big] - f(h_x - y) - e(r_x - y) = 1.\end{aligned}$$

(c) When the rotors from 1 to $y-1$ point inwards ($\downarrow$) and from $y$ to $h_x$ point outwards ($\uparrow$), then $r_x = y - 1$, and we have

$$\begin{aligned}\sigma_m(x, y) = &\frac{1}{2}\big[f(h_x, r_x, y - 1\big) + e(0) - 1\big]\\ &+ \frac{1}{2}f(h_x, r_x, y + 1) - f(h_x, r_x, y) = 1.\end{aligned}$$

(d) The last case which can appear is when all rotors from $y$ to 0 point inwards ($\downarrow$), and from $y+1$ to $h_x$ outwards ($\uparrow$). Then $r_x = y$ and

$$\begin{aligned}\sigma_m(x,y) &= \frac{1}{2}\big[f(h_x, r_x, y-1) + e(1) - 1\big] \\ &\quad + \frac{1}{2} f(h_x, r_x, y+1) - f(h_x, r_x, y) - 1 = 1.\end{aligned}$$

Therefore, for all $(x,y) \in B_m$, with $y \geq 2$, $\sigma_m(x,y) = 1$, which implies that the points $(x,y)$ are left with one particle after firing $u_m$ times.

For the rest of the proof we shift the coordinate system such that the point $(-m,0)$ lies at the origin, which means that we can work with the function $u'$ directly. Since the function $u'$ does not depend on $m$, most of what follows holds independently of $m$. Only to deal with the center point of the set $B_m$ (Case 4), we need to take the parameter $m$ into account.

**Case 2.** $y = 1$: Consider $\sigma_m(z)$ for the vertex $z = (x,1)$. For $x \geq 9$ the number of inwards pointing arrows $r_x$ is always greater than 2. So with the exception of a finite number of exceptional points ($x \in \{1,2,5,8\}$), all relevant rotors on the teeth, are pointing inwards ($\downarrow$) and the vertex $z$ receives $\left\lceil \frac{1}{2} u'(x,2) \right\rceil$ particles from its upper neighbour. For $x \geq 3$, the number $u'(x,0)$ is divisible by 4, so all neighbours of $(x,0)$ receive exactly the same amount of particles. Hence

$$\sigma_m(x,1) = \frac{1}{4} u'(x,0) + \frac{1}{2}\big(u'(x,2) + 1\big) - u'(x,1),$$

if $z$ is non-exceptional. Since the values of $u'$ involved here depend on $r_x$, we need to check each congruence class $x$ mod 3 separately. In all three cases it is an easy computation to show that $\sigma_m(z) = 1$.

At the exceptional points $z = (x,1)$, for $x \in \{1,2,5,8\}$, the correctness of the function $u'$ can be verified directly.

**Case 3.** $x \neq m$ **and** $y = 0$: On the $x$-axis, the points $z = (x,0)$ for $x \in \{0,1,2,3\}$ are again exceptional and need to be checked separately.

For $x \geq 4$, the vertex $z$ receives particles from $(x-1,0)$, $(x+1,0)$, $(x,1)$, $(x,-1)$. Here $u'(x-1,0)$ and $u'(x+1,0)$ are again both divisible by 4. By symmetry $u'(x,1) = u'(x,-1)$, and the number of inward pointing arrows $r_x \geq 1$ in this case, hence $z$ receives $u'(x,1)+1$ particles from its upper and lower neighbours combined. Thus

$$\begin{aligned}\sigma_m(x,0) &= \frac{1}{4}u'(x-1,0) + \frac{1}{4}u'(x+1,0) + u'(x,1) + 1 - u'(x,0) \\ &= \frac{1}{4}\big[2\tilde{u}(h_{x-1}, r_{x-1}, 0) - 2\big] + \frac{1}{4}\big[2\tilde{u}(h_{x+1}, r_{x+1}, 0) - 2\big] \\ &\quad + \tilde{u}(h_x, r_x, 1) - \big[2\tilde{u}(h_x, r_x, 0) - 2\big].\end{aligned} \tag{4.7}$$

Depending on the congruence class mod 3 of $x$, we substitute the corresponding branch of the function $r_x$ in equation (4.7). In all cases $\sigma_m(x,0) = 1$ holds.

**Case 4.** Midpoint $x = m$, $y = 0$: Everything until now was independent on the number of particles $|\mathcal{B}_m|$. Since $u_m$ is created from $u'$ by translation and reflection, the vertex $z = (m, 0)$ after translation corresponds to the origin of the cluster.

At the start of the process, $|\mathcal{B}_m|$ particles are present at $z$, so $\sigma_0(z) = |\mathcal{B}_m|$. We assume that $m$ is big enough, so that none of the neighbours of $z$ is an exceptional point.

By symmetry, $z$ receives $\frac{1}{2}u'(x-1,0)$ particles from its neighbours on the $x$ axis, and $u'(x,1)+1$ particles from its neighbours on the teeth. Hence

$$\sigma_m(z) = \sigma_0(z) + \frac{1}{2}u'(x-1,0) + u'(x,1) + 1 - u'(x,0).$$

Substituting the formulas obtained for $|\mathcal{B}_m|$ mod 3 in (4.3), into the previous equation, gives the desired result $\sigma_m(z) = 1$.

Finally, we need to check that the rotor configuration $\rho_m$ is acyclic. For $m \leq 2$, this follows again directly from Figure 4.3. If $m \geq 3$, we again work with shifted coordinates. It is clear from the previous section that $\rho_m$ restricted to each "tooth" is acyclic. Hence it suffices to check that no cycles are created by rotors on the $x$-axis. If $z = (x, 0)$, the odometer $u_m(z)$ is divisible by 4, except when $x = 2$. So the rotors at these vertices point in the same direction as in the start configuration $\rho_0$. The odometer at the exceptional point $w = \big(2, 0\big)$ is $u'(w) = 23 \equiv 3 \pmod 4$ independent of $m$. Hence, these rotors point in the direction of one "tooth". If the rotor at position $(2, 1)$ points towards the $x$-axis, it creates a creates a directed cycle. By (4.5), we have $r_2 = 0$, which means that all arrows on these "tooth" are pointing outwards. Hence the rotor at $w$ does not close a cycle. See Figure 4.4 for a visualisation of the rotor configurations under consideration.

Therefore all properties of Theorem 4.1.3 are satisfied which proves the statement. $\square$

*Proof of Theorem 4.2.1.* In the case $m \leq 2$, the statement of the Theorem follows by direct calculation of the respective aggregation clusters, see Figure 4.3.

For $m \geq 3$ it follows from the previous Lemma. $\square$

## 4.3 Rotor Weights

In Theorem 4.2.1 we proved a shape result for the rotor-router model for a fixed initial rotor configuration. Similar theorems can be proved for different start configurations, but it would be interesting to prove a shape theorem which holds for arbitrary initial configuration.

Computer simulations, as well as all that is known for rotor-router aggregation on $\mathbb{Z}^d$ and other state spaces, suggest that $R_m$ does not depend on the initial choice of the rotor configuration up to constant fluctuations. In this section we give an inner bound for the cluster $R_m$ which holds for arbitrary initial configuration of the rotors and is independent of the rotor sequence.

The method relies on an idea of Holroyd and Propp [HP10], which they use to show a variety of inequalities concerning rotor-walks and random walks.

Start with a particle distribution $\sigma_0 : G \to \mathbb{N}$ and rotor configuration $\rho_0 : G \to G$ such that $\rho_0(z) = c(z)_0$ for all $z \in G$, that is, all initial rotors point to the first element in the rotor sequence. We further assume that $\sigma_0$ has finite support, i.e., there are only finitely many particles in the system, so that we don't need to deal with questions of convergence. We will route particles in the system, and this gives rise to a sequence $(\rho_t, \sigma_t)_{t\geq 0}$ of particle and rotor configurations at every time $t$. To each of the possible states $(\rho_t, \sigma_t)$ of the system, we will assign a weight.

Fix a function $h : G \to \mathbb{R}$. We define the *Particle weights* at time $t$ to be

$$\mathbf{W_P}(t) = \sum_{z\in G} \sigma_t(z)h(z). \tag{4.8}$$

Define the *Rotor weights* of single points $z \in G$ as

$$w(z,k) = \begin{cases} 0, & \text{for } k = 0 \\ w(z,k-1) + h(z) - h\big(z_{k \bmod d(z)}\big), & \text{for } k > 0, \end{cases} \tag{4.9}$$

where $z_i$ is the $i$-th neighbour of $z$ in the rotor sequence $c(z)$. Notice that, for $k \geq d(z)$,

$$w(z,k) = w\big(z, k - d(z)\big) - d(z) \bigtriangleup h(z). \tag{4.10}$$

The total *Rotor weights* are given by

$$\mathbf{W_R}(t) = \sum_{z \in G} w(z, u_t(z)),$$

where $u_t(z)$ is the odometer function of this process, that is, the number of particles sent out by the vertex $z$ in the first $t$ steps. Note that $\rho_0$ is chosen in such a way that, if $i = u_t(z) \bmod d(z)$, then $z_i = \rho_t(z)$ for all $t \geq 0$ and $z \in G$.

**Lemma 4.3.1.** *The sum of rotor and particle weights $\mathbf{W_P}(t) + \mathbf{W_R}(t)$ is constant.*

*Proof.* We show that $\mathbf{W_P}(t) + \mathbf{W_R}(t) = \mathbf{W_P}(t+1) + \mathbf{W_R}(t+1)$.

Let $z \in G$ be the vertex from which the particle is routed away at time $t$. The particle moves in the direction of the new rotor at $z$, that is, to the vertex $\rho_{t+1}(z)$. Therefore, for the particle weights we have

$$\mathbf{W_P}(t+1) = \mathbf{W_P}(t) - h(z) + h\big(u_{t+1}(z)\big).$$

Since $u_t(w) = u_{t+1}(w)$ for all $w \neq z$, all rotor-weights, except the one at $z$, stay the same. By (4.9) we get

$$\mathbf{W_R}(t+1) = \mathbf{W_R}(t) + h(z) - h\big(u_{t+1}(z)\big).$$

$\square$

In the case of rotor-router aggregation, the initial particle configuration is $\sigma_0 = n \cdot \delta_o$, that is, we start with $n$ particles at the origin and we route a specific particle only if there is at least one other particle at the same position. The process terminates when no two particles are at the same position. By the Abelian property, regardless of the order of particle routings, this process produces the same result as the rotor-router aggregation process we have defined in Section 2.2.2. Write $(\sigma_{\text{end}}, \rho_{\text{end}})$ for the state of the system reached when the rotor-router cluster $R_n$ has been created. By definition

$$\sigma_{\text{end}}(z) = \mathbf{1}_{\{z \in R_n\}}.$$

We use the weight function

$$h(z) = h_y(z) = \frac{G_n(y,z)}{d(z)}, \tag{4.11}$$

where $G_n$ is the Green function stopped upon exiting the set $S_n$ of fully occupied sites for the divisible sandpile with mass distribution $\mu_0 = n \cdot \delta_o$, defined in Theorem 2.2.2.

Note that, for $y \in S_n$,

$$\triangle h_y(z) = \begin{cases} -\frac{1}{d(z)}, & \text{for } y = z \\ 0, & \text{otherwise.} \end{cases} \tag{4.12}$$

The particle weights at time $t = 0$ are given by

$$\mathbf{W_P}(0) = n h_y(o), \tag{4.13}$$

while the rotor weights $\mathbf{W_R}(0) = 0$. At the end of the process, i.e., at time $t = $ end we have

$$\mathbf{W_P}(\text{end}) = \sum_{z \in R_n} h_y(z) \leq \sum_{z \in S_n} h_y(z), \tag{4.14}$$

since $h_y$ is equal to 0 outside of $S_n$. For the rotor weights we get from (4.10)

$$\mathbf{W_R}(\text{end}) = \sum_{z \in R_n} \left\lfloor \frac{u_R(z)}{d(z)} \right\rfloor \big( - d(z) \triangle h_y(z) \big) + \sum_{z \in R_n} w(z, k_z), \tag{4.15}$$

where $u_R$ is the rotor odometer function and $k_z = u_R(z) \bmod d(z)$. Using (4.12) and (4.9) in the previous equation (4.15) we get

$$\begin{aligned} \mathbf{W_R}(\text{end}) &= \left\lfloor \frac{u_R(y)}{d(y)} \right\rfloor + \sum_{z \in R_n} \sum_{i=0}^{k_z} \big( h_y(z) - h_y(z_i) \big) \\ &\leq \frac{u_R(y)}{d(y)} + \sum_{z \in S_n} \sum_{w \sim z} |h_y(z) - h_y(w)|. \end{aligned} \tag{4.16}$$

Since the total weights are invariant under routing of particles, it follows that

$$\mathbf{W_P}(0) + \mathbf{W_R}(0) = \mathbf{W_P}(\text{end}) + \mathbf{W_R}(\text{end}). \tag{4.17}$$

From (4.17), together with (4.13), (4.14) and (4.16), we obtain

$$\sum_{z \in S_n} \big( n\delta_0(z) - 1 \big) h_y(z) \leq \frac{u_R(y)}{d(y)} + \sum_{z \in S_n} \sum_{w \sim z} |h_y(z) - h_y(w)|. \tag{4.18}$$

If we write $v(y)$ for the left hand side of inequality (4.18), that is

$$v(y) = \sum_{z \in S_n} \big( n\delta_0(z) - 1 \big) h_y(z),$$

then $v(y) = 0$ for $y \notin S_n$, and because of the linearity of the Laplacian, we get

$$\triangle v(y) = \frac{1}{d(y)} \big( 1 - n\delta_0(z) \big) \text{ for } y \in S_n.$$

By (3.4), the normalized odometer function $u_n$ of the divisible sandpile on $G$ with initial mass distribution $\mu_0 = n \cdot \delta_o$ satisfies exactly the same Dirichlet problem, hence $v(y) = u_n(y)$ and we get the following result, which compares the odometer function of rotor-router aggregation with the odometer function of the divisible sandpile.

**Proposition 4.3.2.** *Let $u_n$ be the normalized odometer function of the divisible sandpile with initial mass distribution $\mu_0 = n \cdot \delta_o$, and $u_R$ the odometer function of rotor-router aggregation with $n$ particles starting at the origin $o \in G$. Then, for all $y \in G$,*

$$u_n(y) \leq \frac{u_R(y)}{d(y)} + \mathbf{W}_{\text{rest}}(y), \tag{4.19}$$

*with*

$$\mathbf{W}_{\text{rest}}(y) = \sum_{z \in S_n} \sum_{w \sim z} \left| \frac{G_n(y,z)}{d(z)} - \frac{G_n(y,w)}{d(w)} \right|. \tag{4.20}$$

Inequality (4.19) has been derived by Levine and Peres [LP09] in the case of $\mathbb{Z}^d$ using a different method. In Section 4.4 we use a variant of their approach to prove an inner bound of the rotor-router cluster, which holds for arbitrary regular graphs.

For trees, the upper bound of the rotor weights $\mathbf{W}_{\text{rest}}(y)$ can be written in terms of the estimated distance of the starting point of a random walk to the point where it first exits $S_n$.

**Proposition 4.3.3.** *If $G$ is a tree and $\mathsf{d}(\cdot,\cdot)$ is the graph distance on $G$, then*

$$\mathbf{W}_{\text{rest}}(x) = 2\mathbb{E}_x\big[\mathsf{d}(x, X_T)\big] - 2,$$

*where $T = \inf\big\{t \geq 0 : X_t \notin S_n\big\}$, and $X_t$ is the simple random walk on $G$.*

*Proof.* For $y \sim z$ let $N_{yz}$ be the number of transitions from $y$ to $z$ before the random walk exits $S_n$. Then

$$\mathbb{E}_x\big[N_{yz} - N_{zy}\big] = \frac{G_n(x,y)}{d(y)} - \frac{G_n(x,z)}{d(z)}.$$

See also [LP, Proposition 2.2] for more details.

Since $G$ is a tree, the net number of crossings of each edge is smaller or equal to one, i.e.,

$$\big|\mathbb{E}_x\big[N_{yz} - N_{zy}\big]\big| \leq 1.$$

We consider $G$ as a tree rooted at $x$, and denote by $\pi_{x,z}$ the shortest path from $x$ to $z$. For $y \neq x$, write $y^-$ for the parent of $y$, i.e., the unique neighbour of $y$ that lies on the shortest path $\pi_{x,y}$. With this notation we get

$$\begin{aligned}\sum_{\substack{y,z\in S_n\\ y\sim z}} \left|\frac{G_n(x,y)}{d(y)} - \frac{G_n(x,z)}{d(z)}\right| &= \sum_{\substack{y,z\in S_n\\ y\sim z}} \left|\mathbb{E}_x\big[N_{yz} - N_{zy}\big]\right| \\ &= 2\sum_{\substack{y\in S_n\\ y\neq x}} \mathbb{E}_x\big[N_{y^-y} - N_{yy^-}\big],\end{aligned}$$

where the last equality is due to the antisymmetry of $N_{yz} - N_{zy}$. Let

$$C_y = \big\{z \in S_n : \ y \in \pi_{x,z}\big\}$$

be the cone of $y$. The random variable $N_{y^-y} - N_{yy^-}$ is either zero or one, the latter if the random walk exits $S_n$ in the cone $C_y$, hence

$$\begin{aligned}\mathbf{W}_{\text{rest}}(x) &= 2\sum_{\substack{y\in S_n\\ y\neq x}} \mathbb{P}_x\big[X_T \in C_y\big] \\ &= 2\sum_{\substack{y\in S_n\\ y\neq x}}\sum_{z\in C_y} \mathbb{P}_x\big[X_T = z\big]\end{aligned}$$

For all $z \in \partial S_n$ we have $\#\big\{y \in S_n \setminus \{x\} : \ z \in C_y\big\} = \mathsf{d}(x,z) - 1$, therefore

$$\begin{aligned}\mathbf{W}_{\text{rest}}(x) &= 2\sum_{z\in\partial S_n} \mathbb{P}_x\big[X_T = z\big]\big(\mathsf{d}(x,z) - 1\big) \\ &= 2\mathbb{E}_x\big[\mathsf{d}(x,X_T)\big] - 2.\end{aligned}$$

□

By Proposition 4.3.2 and 4.3.3 and Corollary 3.5.3, one needs an upper bound for the expected distance to the exit point of a random walk, in order to derive an inner estimate of the rotor-router cluster.

Using the the trivial upper estimate

$$\begin{aligned}\mathbb{E}_z\big[\mathsf{d}(z,X_T)\big] &\leq \max\big\{\mathsf{d}(z,w) : w \in \partial S_n\big\} \\ &= |x| + |y| + ln^{2/3},\end{aligned} \tag{4.21}$$

with $z = (x,y)$ and $l = \frac{1}{2}\big(\frac{3}{2}\big)^{1/3}$ as in Theorem 2.2.2, we can show the following weak inner bound.

**Theorem 4.3.4.** *Let $R_n$ be the rotor-router cluster of $n$ particles on the comb. Then, for $n \geq n_0$ and independently of the initial rotor configuration and the choice of rotor sequence, we have the following inner bound*

$$\tilde{\mathcal{B}}_n \subset R_n,$$

*where*

$$\tilde{\mathcal{B}}_n = \Big\{(x,y) \in \mathcal{C}_2 : \ |x| \leq kn^{1/3} - c_1 n^{1/6}, \\ |y| \leq l\left(n^{1/3} - \frac{x}{k}\right)^2 + c_2 x - c_3 n^{1/3}\Big\},$$

*$k = \left(\frac{3}{2}\right)^{2/3}$ and $l = \frac{1}{2}\left(\frac{3}{2}\right)^{1/3}$ and $c_1$, $c_2$ and $c_3$ are constants.*

*Proof.* By definition of the odometer function $u_R$

$$\big\{z \in \mathcal{C}_2 : u_R(z) > 0\big\} \subset R_n,$$

and by Proposition 4.3.2 together with Proposition 4.3.3, we have for vertices $z = (x, y)$

$$\begin{aligned}\frac{u_R(z)}{d(z)} &\geq u_n(z) - 2\mathbb{E}_z\big[\mathsf{d}(z, X_T)\big] + 2 \\ &\geq u_n(z) - 2\big(|x| + |y| + ln^{2/3}\big) + 2,\end{aligned}$$

where the last inequality is due to (4.21). By Corollary 3.5.3, we have a lower bound of the sandpile odometer $u_n$ for $z \in S_n$

$$\gamma_n(z) - a \leq u_n(z),$$

where $a$ is a positive constant smaller than 2, and $\gamma_n$ is the function defined in (3.15).

Thus, to derive an inner bound, it suffices to check for which $z = (x, y) \in S_n$ the inequality

$$\gamma_n(x,y) - 2\big(|x| + |y| + ln^{2/3}\big) > 0 \tag{4.22}$$

holds. By symmetry it is enough to consider $x, y \geq 0$.

We first check inequality (4.22) on a "tooth" of the comb, that is, for a fixed $x$. The function $\gamma_n$ is given as

$$\gamma_n(x,y) = \frac{1}{2}\left(y - \frac{n_x}{2}\right)^2,$$

where $n_x$ is the amount of mass that ends up in the $x$-"tooth" of the sandpile. Since $x$ is fixed, we can treat $n_x$ as a constant. Hence the right hand side of (4.22) is a quadradic polynomial in $y$ with smallest root

$$y_x = 2 + \frac{n_x}{2} - \sqrt{4 + \frac{k}{l}n^{2/3} + 2n_x + 4x}.$$

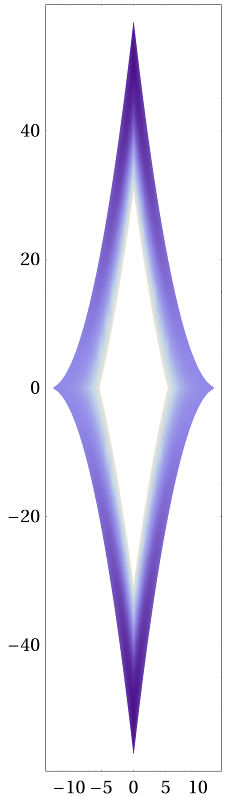



Figure 4.5: Picture of the inner estimate of $R_n$ in comparison to sandpile cluster $S_n$, for $n = 1000$. The white area is the area where the inequality (4.22) is valid, and corresponds to the set $\tilde{\mathcal{B}}_n$ of Theorem 4.3.4. The coloring is based on the value of the right-hand side of (4.22).

Substituting $n_x$ as calculated in (3.12), and expansion around $n = \infty$ gives

$$y_x = ln^{2/3} - \frac{1}{2l}n^{1/3}x + \frac{x^2}{3} + \frac{2+\sqrt{6}}{3}x + c_1 n^{1/3} - c_2\frac{x^4}{n}.$$

Since $(x, y) \in S_n$, we have the bound $x \leq kn^{1/3}$, hence

$$y_x = l\left(n^{1/3} - \frac{x}{k}\right)^2 + \frac{2+\sqrt{6}}{3}x - cn^{1/3}, \tag{4.23}$$

for $n \geq n_0$, and a positive constant $c$.

To get a bound on the $x$-axis, we calculate for which $x > 0$ the inequality $y_x > 0$ is satisfied. Since $y_x$ is a polynomial of degree 2 in $x$ this is easy to do, and again by series expansion around $n = \infty$ we obtain

$$x \leq k \cdot n^{1/3} - c_3 n^{1/3}, \tag{4.24}$$

for $n \geq n_0$. The inner bound for $R_n$ now follows from (4.24) together with (4.23). □

Figure 4.5 shows the inner estimate of the rotor-router cluster that was proved in Theorem 4.3.4. The estimate could be improved if one had a substantially better upper bound for $\mathbb{E}_z[\mathsf{d}(z, X_T)]$. For this a good understanding of the harmonic measure of the set $\mathcal{B}_n$, that is, the probability $\mathbb{P}_z\big[X_z = y\big]$, will probably be essential. We will be investigate the harmonic measure in Chapter 5.

## 4.4 A Universal Inner Bound for Rotor-Router Aggregation

In this Section a universal inner estimate for rotor-router aggregation is proved, which relates the rotor-router cluster to a divisible sandpile cluster with a smaller mass.

Like before, let $G$ be a locally finite directed graph, and $p(x, y)$ the transition probabilities of the simple random walk on $G$. Let $X_t$ be the trajectory of the random walk. The neighbourhood relation on $G$ is denoted by $x \sim y$. We will think of each edge as a pair of directed edges $x \to y$ and $y \to x$.

For a function $f : G \to \mathbb{R}$ and $x \sim y$, define

$$\nabla f(x, y) = f(y) - f(x).$$

If $s$ is a function on the directed edges of $G$, that is, $s : G \times G \to \mathbb{R}$, the *divergence* of $s$ is defined as

$$\operatorname{div} s(x) = \sum_{y \sim x} p(x, y) s(x, y).$$

Finally the Laplace operator can be written as

$$\triangle f(x) = \operatorname{div}(\nabla f)(x) = \sum_{y \sim x} p(x, y) f(y) - f(x). \tag{4.25}$$

As in Section 4.3, let $u_R(x)$ be the odometer function of rotor-router aggregation with $n$ particles starting at a chosen root vertex $o$. Recall that $u_R(x)$ is the total number of particles that are sent out from vertex $x$ during the whole aggregation process. For ease of notation we will also use the normalized odometer function denoted by $u'_R(x) = \frac{u_R(x)}{d(x)}$.

Let $N(x, y)$ be the number of particles routed along the edge $x \to y$, and $N(x, y) = 0$ if $y$ is not a neighbour of $x$. Additionally let

$$K(x, y) = N(x, y) - N(y, x).$$

The number of edge crossings $N(x,y)$ can be bounded by the odometer function at $x$. If $y$ is a neighbour of $x$, then

$$\frac{u_R(x)-d(x)+1}{d(x)} \le N(x,y) \le \frac{u_R(x)+d(x)-1}{d(x)},$$

which can be written as

$$u_R'(x)-1+\frac{1}{d(x)} \le N(x,y) \le u_R'(x)+1-\frac{1}{d(x)}. \tag{4.26}$$

Substracting inequalities (4.26) for a pair of edges $x \to y$ and $y \to x$ gives

$$-2+\frac{1}{d(x)}+\frac{1}{d(y)} \le \nabla u_R'(x,y)+K(x,y) \le 2-\frac{1}{d(x)}-\frac{1}{d(y)}.$$

Thus, for some antisymmetric function $R(x,y)$, with

$$|R(x,y)| \le 2-\frac{1}{d(x)}-\frac{1}{d(y)}, \tag{4.27}$$

we have

$$\nabla u_R'(x,y) = -K(x,y)+R(x,y). \tag{4.28}$$

Using (4.25), the Laplacian of the normalized odometer function $u_R'$ can be written in terms of $K$ and $R$

$$\begin{aligned} \triangle u_R'(x) &= \operatorname{div}(\nabla u_R')(x) = \sum_{y\sim x} p(x,y)\nabla u_R'(x,y) \\ &= \frac{1}{d(x)}\sum_{y\sim x}\big(-K(x,y)+R(x,y)\big), \end{aligned} \tag{4.29}$$

where the last equality is due to (4.28).

Note that, for each $x \neq o$, the number of particles that arrive at $x$ is at most one more than the number of particles that are sent out from $x$, hence

$$\sum_{y\sim x} K(x,y) \ge -1.$$

If $x = o$, one has to take into account that all $n$ particles start at $o$, which gives

$$\sum_{y\sim 0} K(0,y) = n-1.$$

This together with (4.29) implies the following upper bound for the Laplacian of the odometer function in terms of $R$

$$\triangle u_R'(x) \le \frac{1}{d(x)}\big(1-n\mathbf{1}_{\{x=0\}}\big)+\operatorname{div} R(x). \tag{4.30}$$

Now let $B$ be a subset of the vertex set of $G$ which contains the root $o$, and

$$T = \inf\big\{t \geq 0 : \ X_t \notin B\big\}$$

be the stopping time of the first exit of $B$.

By telescoping we have

$$\begin{aligned} u'_R(x) - \mathbb{E}_x\big[u'_R(X_T)\big] &= \sum_{k\geq 0} \mathbb{E}_x\big[u'_R(X_{k\wedge T}) - u'_R(X_{(k+1)\wedge T})\big] \\ &= \sum_{k\geq 0} \mathbb{E}_x\big[ - \triangle u'_R(X_k)\mathbf{1}_{\{k<T\}}\big], \end{aligned}$$

which, by (4.30), is bigger than

$$\begin{aligned} &\geq \sum_{k\geq 0} \mathbb{E}_x \left[\left(\frac{n}{d(X_k)}\mathbf{1}_{\{X_k=0\}} - \frac{1}{d(X_k)} - \operatorname{div} R(X_k)\right)\mathbf{1}_{\{k<T\}}\right] \\ &= nG_B(x,0) - \mathbb{E}_x\left[\sum_{k=0}^{T-1} \frac{1}{d(X_k)}\right] - \sum_{k\geq 0} \mathbb{E}_x\left[\operatorname{div} R(X_k)\mathbf{1}_{\{k<T\}}\right]. \end{aligned}$$

Recall that $G_B(x,y) = \mathbb{E}_x\big[\sum_{k=0}^{T-1} \mathbf{1}_{\{X_k=y\}}\big]$ is the Green function stopped at the first exit of $B$. If we define the function $f(x)$ as

$$f(x) = \mathbb{E}_x\big[u'_R(X_T)\big] + nG_B(x,0) - \mathbb{E}_x\left[\sum_{k=0}^{T-1} \frac{1}{d(X_k)}\right],$$

we get the inequality

$$u'_R(x) - f(x) \geq -\sum_{k\geq 0} \mathbb{E}_x\left[\operatorname{div} R(X_k)\mathbf{1}_{\{k<T\}}\right]. \tag{4.31}$$

An equivalent relation to (4.31) has been derived by Levine and Peres in [LP09] in the case of $\mathbb{Z}^d$. If one makes use of the antisymmetry of $R(x,y)$, from (4.31) one can derive inequality (4.19), of Section 4.3, with an additional small error term.

We can further rewrite the right hand side of (4.31) as:

$$\begin{aligned} \sum_{k\geq 0} \mathbb{E}_x\left[\operatorname{div} R(X_k)\mathbf{1}_{\{k<T\}}\right] &= \sum_{k\geq 0} \mathbb{E}_x\left[\sum_{z\sim X_k} \frac{1}{d(X_k)} R(X_k,z)\mathbf{1}_{\{k<T\}}\right] \\ &= \sum_{k\geq 0} \mathbb{E}_x\left[\sum_{y\in B} \mathbf{1}_{\{k<T\}}\mathbf{1}_{\{X_k=y\}}\frac{1}{d(y)}\sum_{z\sim y} R(y,z)\right] \end{aligned}$$

Since $X_T \notin B$, we can write the event $[k < T] \cap [X_k = y]$ as $[X_{k\wedge T} = y]$. Evaluation of the expectation now gives

$$
\begin{aligned}
&= \sum_{k\geq 0}\sum_{y\in B} \mathbb{P}_x\left[X_{k\wedge T} = y\right] \frac{1}{d(y)} \sum_{z\sim y} R(y,z) \\
&\leq \sum_{y\in B} G_B(x,y) \frac{1}{d(y)} \sum_{z\sim y} \left(2 - \frac{1}{d(y)} - \frac{1}{d(z)}\right),
\end{aligned}
$$

where in the last inequaltity we changed the order of summation and applied the upper bound (4.27) for $|R(x,y)|$. Hence

$$
\begin{aligned}
\sum_{k\geq 0} \mathbb{E}_x\left[\operatorname{div} R(X_k)\mathbf{1}_{\{k<T\}}\right] &\leq \sum_{y\in B} G_B(x,y)\left[2 - \frac{1}{d(y)} - \frac{1}{d(y)}\sum_{z\sim y}\frac{1}{d(z)}\right] \\
&= 2\mathbb{E}_x[T] - \mathbb{E}_x\left[\sum_{k=0}^{T-1}\frac{1}{d(X_k)}\right] - \sum_{y\in B}\frac{G_B(x,y)}{d(y)}\sum_{z\sim y}\frac{1}{d(z)}.
\end{aligned}
$$

If we plug this into (4.31) we get some cancellation, and obtain the following lower bound for the normalized rotor odometer function

$$u_R(x) \geq \tilde{f}(x), \tag{4.32}$$

where the function $\tilde{f}(x)$ is defined as

$$
\begin{aligned}
\tilde{f}(x) &= \mathbb{E}_x\left[u'_R(X_T)\right] + nG_B(x,0) + 2\mathbb{E}_x[T] - \sum_{y\in B}\frac{G_B(x,y)}{d(y)}\sum_{z\sim y}\frac{1}{d(z)} \\
&= \mathbb{E}_x\left[u'_R(X_T)\right] + \sum_{y\in B}\left(n\cdot\mathbf{1}_{\{y=0\}} - 2d(y) + \sum_{z\sim y}\frac{1}{d(z)}\right)\frac{G_B(x,y)}{d(y)}.
\end{aligned}
$$

The function $\mathbb{E}_x\left[u'_R(X_T)\right]$ is harmonic for $x \in B$ and, if we set $g_y(x) = \frac{G_B(x,y)}{d(y)}$, we obtain

$$
\triangle g_y(x) = \begin{cases} -\frac{1}{d(x)}, & \text{if } y = x \\ 0, & \text{otherwise,} \end{cases}
$$

for $x \in B$. Hence, the Laplacian of $\tilde{f}(x)$ for $x \in B$ is equal to

$$\triangle \tilde{f}(x) = \frac{1}{d(x)}\left(2d(x) - \sum_{z\sim x}\frac{1}{d(z)} - n\cdot\mathbf{1}_{\{x=0\}}\right). \tag{4.33}$$

Additionally $\tilde{f}(x) \geq 0$ for $x \in \partial B$. If we compare this with (3.4), it follows by the Minimum Principle for superharmonic functions, that $\tilde{f}(x)$ is bigger than the odometer function of a divisible sandpile starting with mass $n$ at the origin, where the depth $h(x)$ of the holes depends on the site and is given by

$$h(x) = 2d(x) - \sum_{z \sim x} \frac{1}{d(z)},$$

where the set $B$ is chosen as the sandpile cluster of this process. It then follows from (4.32) that $B$ is a subset of the rotor-router cluster $R(n)$.

In the special case when $G$ is a regular graph with degree $d$, we have

$$h(x) = 2d - 1,$$

and (4.33) simplifies to

$$\triangle \tilde{f}(x) = \frac{2d-1}{d} \left( 1 - \frac{n}{2d-1} \cdot \mathbf{1}_{\{x=0\}} \right),$$

which implies the following.

**Proposition 4.4.1.** *Let $G$ be a regular graph with degree $d$ and root vertex $o$, and let $R(n)$ be the rotor-router cluster of $n$ particles starting on the root $o$. Further, let $S(n)$ be the divisible sandpile cluster for the mass distribution $\mu_0(x) = n \cdot \delta_o(x)$. Then*

$$S\left(\tfrac{n}{2d-1}\right) \subset R(n).$$

Chapter 5

# The Harmonic Measure

In this chapter we give an application of rotor-router walks in the calculation of harmonic measures of random walks, that is, the hitting distribution of a finite set. While in priciple this approach allows to obtain exact results, as we will show for a few examples, it is difficult to apply in concrete cases as it requires exact knowledge of the odometer function of the rotor-router walk, and at least some insight in the structure of the *Abelian sandpile group* of the set under consideration. The connection of the *Abelian sandpile group* to the rotor router model has been established in the physics literature; see [PPS98, PDDK96]. One can define a group based on the action of a particle which performs a rotor-walk on the rotor configuration. This *rotor-router group* is Abelian and isomorphic to the Abelian sandpile group. This isomporphism has been proven formally in [LL09]. For a self-contained introduction see the overview paper of HOLROYD, LEVINE, ET.AL. [HLM+08].

As before, let $G$ be a locally finite, connected graph, and $B$ be a finite subset of vertices $G$. Write

$$\partial_I B = \{x \in B : \exists y \notin B \text{ with } x \sim y\}$$

for the inner boundary of $B$, and we will write $B^\circ = B \setminus \partial_I B$. The vertices of $\partial_I B$ will also be called the *sink*.

Similarly to Definition 4.1.1 of Chapter 4, we define the *particle addition operator* $E_v$, for each vertex $v \in B^\circ$. For a rotor configuration $\rho$, let

$$E_v(\rho) = \rho',$$

where $\rho'$ is the rotor configuration obtained from $\rho$ by adding a new particle at vertex $v$, and letting it perform a rotor-router walk until the particle reaches a sink vertex in $\partial_I B$ for the first time.

By the abelian property of rotor-router walks the operators $E_v$ commute, and they can be used to a define an abelian group, see [HLM+08] for details and [HLM+08, Lemma 3.10] for the proof of the following statement.

**Lemma 5.0.2.** *The particle addition operator $E_v$ is a permutation on the set of acyclic rotor configurations on $B^\circ$.*

The *rotor-router group* of $B^\circ$ is defined as the subgroup of permutations of oriented spanning trees rooted at the sink (that is, acyclic rotor configurations) generated by $\{E_v : v \in B^\circ\}$. For every finite graph $B^\circ$ the rotor-router group is a finite abelian group, which is isomorphic to the *abelian sandpile group*. See again [HLM+08] for details.

Now, consider again the stopping time

$$T = \inf\{t \geq 0 : X_t \in \partial_I B\}.$$

Let $\nu_x(z) = \mathbb{P}_x[X_T = z]$ be the harmonic measure of $B$, with starting point $x$. If the starting point is the origin $o$, we will drop the subscript and write $\nu(z) = \nu_o(z)$.

The harmonic measure $\nu(z)$ is important for the outer estimate for both IDLA and rotor router aggregation. We can use rotor weights as in Section 4.3 to calculate the harmonic measure for subsets of the comb, if we use the harmonic measure itself as the weight function $h(x)$, which is used in (4.8) and (4.9) to define the particle weights $\mathbf{W_P}(t)$ and rotor weights $\mathbf{W_R}(t)$. Fix a vertex $z \in \partial_I B$ and define the weight function as

$$h(x) = h_z(x) = \nu_x(z).$$

Consider the following process. Start with $n$ particles at the origin $o$, and an arbitrary acyclic rotor configuration $\rho_0$. We let the particles perform rotor walks until they reach a vertex in $\partial_I B$ for the first time, where they stop. For each $w \in \partial_I B$, denote by $e(w)$ the number of particles that are in $w$ at the end of this procedure. For all $w \in B$, denote by $u(w)$ the normalized rotor odometer function of this process

$$u(w) = \frac{\text{Number of particles sent out by } z}{d(w)}.$$

Using the invariance of the sum of rotor- and particle-weights under rotor-router walks, as in (4.17), we get

$$nh(o) = \sum_{w \in \partial_I B} e(w)h(w) + \mathbf{W_R}(\text{end}),$$

which reduces to

$$nh(o) = e(z) + \mathbf{W_R}(\text{end}), \tag{5.1}$$

since $h(w) = \nu_w(z) = \delta_w(z)$, if $w \in \partial_I B$.

Since the initial rotor configuration $\rho_0$ is chosen to be acyclic, there exists a number $n$ such that, after all $n$ particles performed their walks, all rotors in $B^\circ$ made only full turns, i.e. $\rho_0 = \rho_{\text{end}}$. Hence, $n$ is the order of $E_o$ in the rotor-router group.

Since $h$ is harmonic on $B \setminus \partial_I B$, using a $n$ with the above property gives $\mathbf{W_R}(\text{end}) = 0$. This together with (5.1) leads to the following equation

$$n \cdot \nu_o(z) = e(z). \tag{5.2}$$

Hence, the harmonic measure $\nu_o$ is proportional to the number of rotor-router particles at the vertices of the boundary.

While a number $n$ with the right property is difficult to calculate, we can still use equation (5.2) in order to derive asymptotics of the harmonic measure of subsets of the comb, and in some cases even to calculate it explicitly.

For this, we consider sets $B_m$ of the type defined in Theorem 4.2.1. For some positive function $r : \mathbb{N}_0 \to \mathbb{N}_0$, define

$$B_m = \big\{(x, y) \in \mathcal{C}_2 : \ |x| \le m, |y| \le r(m - |x|)\big\} \quad \text{for } m \in \mathbb{N}. \tag{5.3}$$

By symmetry of the set $B_m$, it is clear that also $e(w)$ and $\nu(w)$ are symmetric. More precisely, if $w = (w_x, w_y)$ and $w' = (|w_x|, |w_y|)$ then

$$e(w) = e(w') \qquad \text{and} \qquad \nu(w) = \nu(w').$$

Since $e$ and $\nu$ are only defined at the boundary $\partial_I B_m$ we will, for simplicity of notation, write $e(x) = e(x, r(x))$ and $\nu(x) = \nu(x, r(x))$. Additionally, like in the previous section, we will shift the coordinate system by $(m, 0)$ such that the leftmost point of the set $B_m$ has the coordinate $(0, 0)$.

Since all rotors make only full turns, the odometer function $u(w)$ is harmonic outside the origin

$$\triangle u(w) = \begin{cases} 0, & w \in B \setminus \big(\partial_I B \cup \{o\}\big) \\ -n, & w = o, \end{cases} \tag{5.4}$$

and $u(w) = 0$, for $w \in \partial_I B$. Solving the Dirichlet problem (5.4) on the teeth of the comb, gives for $w = (x, y) \in B_m$,

$$u(w) = u(x, y) = e(x) \cdot \big(r(x) - y\big). \tag{5.5}$$

On the $x$-axis, for $x \neq o$, the harmonicity gives

$$u(x + 1, 0) + u(x - 1, 0) + 2u(x, 1) = 4u(x, 0),$$

which together with (5.5) leads to the following recursion for $e(x)$

$$e(x + 1)r(x + 1) + e(x - 1)r(x - 1) - 2e(x)\big(r(x) + 1\big) = 0. \tag{5.6}$$

**Example 5.0.3.** *As a first example where one can calculate the harmonic measure explicitly, consider the function* $r(x) = x^2$*, and its associated set* $B_m$ *as defined in* (5.3)*. From* (5.6) *we get the recursion*

$$e(x+1)(x+1)^2 + e(x-1)(x-1)^2 - 2e(x)(x^2+1) = 0,$$

*for* $m > x > 0$*. Using the special structure of the set* $B_m$*, it is easy to get initial values for the sequence* $e(x)$*. Since* $r(0) = 0$ *and* $r(1) = 1$*, the vertex* $w = (1,0)$ *has three of its neighbours, namely* $(0,0)$*,* $(1,1)$ *and* $(1,-1)$ *on the boundary of* $B_m$*. By construction, the rotor at* $w$ *makes a number of full turns, hence all of these three points receive an equal number of particles from* $w$*. This means that* $e(0) = e(1)$*.*

*By induction, it is easy to see that the sequence* $e(x)$ *is constant. Assuming* $e(x-1) = e(x)$*, the above recursion reduces to*

$$e(x+1)(x+1)^2 - e(x)(x+1)^2 = 0.$$

*Since* $(x+1)^2 > 0$ *for all positive* $x$*, this implies that* $e(x+1) = e(x)$*.*

As $e(x)$ is by construction proportional to the harmonic measure, we get the following.

**Lemma 5.0.4.** *Let* $r(x) = x^2$ *and* $B_m \subset \mathcal{C}_2$ *with*

$$B_m = \big\{(x,y) \in \mathcal{C}_2 : \ |x| \le m, |y| \le r(m-x)\big\} \quad \text{for } m \in \mathbb{N}.$$

*Then, for all* $m \geq 0$*, the harmonic measure* $\nu_0$ *is the uniform measure on* $\partial_I B_m$*.*

**Example 5.0.5.** *A second example where the harmonic measure can be deduced explicitly is the case of a box with*

$$r(x) = \begin{cases} r, & x > 0 \\ 0, & x = 0. \end{cases}$$

*Since the vertex* $(0,0)$ *receives all its particles from vertex* $(1,0)$*, the number of stopped particles in* $(0,0)$ *is the normalized odometer function in* $(1,0)$

$$e(0) = u(1,0) = e(1) \cdot r(1).$$

*Equation* (5.6) *with* $x = 1$ *gives*

$$e(2) = 2e(1)\frac{r+1}{r}, \tag{5.7}$$

*and, for* $x > 1$*, we get as the general solution of the linear recursion*

$$e(x) = \left(\frac{r+1-\sqrt{2r+1}}{r}\right)^x \cdot c_1 + \left(\frac{r+1+\sqrt{2r+1}}{r}\right)^x \cdot c_2.$$

*From equation (5.7) we deduce $c_2 = -c_1$. This gives for the harmonic measure*

$$\nu(x) = \begin{cases} \left[ \left( \frac{r+1+\sqrt{2r+1}}{r} \right)^x - \left( \frac{r+1-\sqrt{2r+1}}{r} \right)^x \right] c, & \text{for } 0 < x \le m \\ 2\sqrt{2r+1}c, & \text{for } x = 0, \end{cases}$$

*where $c$ is a positive constant depending on the width $m$ and height $r$ of the box.*

Much more interesting is in our context the harmonic measure of the fully symmetric rotor router clusters of Theorem 4.2.1. Since $r(x) = \left\lfloor \frac{(x+1)^2}{3} \right\rfloor$ in that case, we have to solve a linear recurrence with non-polynomial coefficients. While an explicit answer is not feasible, we can derive asymptotics of the solution, by converting the recurrence into an equvialent system of linear differential equations. The next section gives a short overview of *singularity analysis* for linear differential equations, which can be used to compute such asympotics.

## 5.1 Singularity Analysis of Linear Differential Equations

In this section we are interested in the asympotics of the coefficients of analytic solutions of ordinary linear differential equations. The presentation follows the book of Flajolet and Sedgewick [FS09].

Let $f(z) = \sum_{n\ge0} f_n z^n$ be a formal power series, then we denote by $[z^n]$ the *coefficient extraction operator*, with definition

$$[z^n]f(z) = f_n.$$

Consider a ordinary linear differential equation of the form

$$c_0(z)\frac{\partial^r}{\partial z^r}f(z) + c_1(z)\frac{\partial^{r-1}}{\partial z^{r-1}}f(z) + \cdots + c_r(z)f(z) = 0, \tag{5.8}$$

where all coefficients $c_j(z)$ are analytic in a simply connected domain $\Omega$. By classical theory, singularities of the solutions of this ODE can only occur at roots $\xi$ of the leading coefficient $c_0$.

From now on we will work with normalized differential equations

$$\frac{\partial^r}{\partial z^r}f(z) + d_1(z)\frac{\partial^{r-1}}{\partial z^{r-1}}f(z) + \cdots + d_r(z)f(z) = 0, \tag{5.9}$$

with $d_j(z) = \frac{c_j(z)}{c_0(z)}$. The $d_j(z)$ are meromorphic functions in the domain $\Omega$, and we define $\omega_\xi(j)$ to be the order of the pole of $d_j$ at the point $\xi$.

**Definition 5.1.1.**

- *The differential equation* (5.9) *has a* singular point *at* $\xi$, *if there exists an index* $1 \leq i \leq r$ *with*
$$\omega_\xi(i) > 0.$$

- *A singularity* $\xi$ *is called* regular, *if*
$$\omega_\xi(1) \leq 1, \quad \omega_\xi(2) \leq 2, \quad \ldots, \quad \omega_\xi(r) \leq r.$$

**Definition 5.1.2.** *For an ODE of the form* (5.9) *and a regular singularity* $\xi$, *the* indicial polynomial $I_\xi(\theta)$ *is defined as*

$$I_\xi(\theta) = \theta^{\underline{r}} + \delta_1 \theta^{\underline{r-1}} + \cdots + \delta_r,$$

*where* $\delta_j = \lim_{z \to \xi} (z - \xi)^j d_j(z)$ *and* $\theta^{\underline{l}} = \theta(\theta - 1) \cdots (\theta - l + 1)$ *denotes the lower factorial.*

We have now collected all preliminaries to be able to state the theorem. For more details see once more Flajolet and Sedgewick [FS09, Theorem VII.10].

**Theorem 5.1.3** (Coefficient asympototics for linear ODE)**.** *Let* $f(z)$ *be analytic at* 0 *and satisfy an ODE of the form* (5.9), *where the coefficients* $d_j(z)$ *are analytic in* $0 < |z| < \rho_1$, *except for a pole at some* $\xi$ *with* $0 < |\xi| < \rho_1$. *If* $\xi$ *is a regular singular point and no two roots of the indicial equation* $I_\xi(\theta) = 0$ *differ by an integer, then there exist constants* $\lambda_1, \ldots, \lambda_r \in \mathcal{C}_2$ *such that, for all* $\rho_0$ *with* $|\xi| < \rho_0 < \rho_1$, *we have*

$$[z^n] f(z) = \sum_{j=1}^{r} \delta_j \Delta_j(n) + O(\rho_0^{-n}).$$

*The* $\Delta_j(n)$ *are of the asymptotic form*

$$\Delta_j(n) \sim \frac{n^{-\theta_j - 1}}{\Gamma(-\theta_j)} \xi^{-n} \left[ 1 + \sum_{k=1}^{\infty} \frac{s_{i,j}}{n^i} \right],$$

*where* $\theta_j$ *are the roots of the indicial equation* $I_\xi(\theta) = 0$.

Note that the $d_j(z)$ can have a singularity of any kind at 0, (see Note VII.47 in [FS09]).

**Remark 5.1.4.** *We will not be able to apply Theorem 5.1.3 directly in the case which is of interest for us, because the requirement that the indicial equation has no two roots that differ by an integer will not be satisfied. In this case one gets additional logarithmic terms and the $\Delta_j(n)$ are asymptotic of the form*

$$\xi^{-n} n^{\beta} \log^{l} n, \tag{5.10}$$

*where $\xi$ is a regular singular point, $\beta$ is a algebraic number satisfying*

$$I_{\xi}(-\beta - 1) = 0,$$

*and $l$ is an integer. For more details see [FS09, page 521].*

## 5.2 Harmonic Measure of $B_m$

We can now state the result about the harmonic measure of the set $B_m$ from Theorem 4.2.1. Using the notation introduced at the beginning of this Chapter, we have the following.

**Lemma 5.2.1.** *Let $r(x) = \left\lfloor \frac{(x+1)^2}{3} \right\rfloor$ and $B_m \subset \mathcal{C}_2$ with*

$$B_m = \big\{(x, y) \in \mathcal{C}_2 : \ |x| \le m, |y| \le r(m - x)\big\} \quad \text{for } m \in \mathbb{N}.$$

*Then for, all $m \ge 0$, the harmonic measure $\nu_o(x)$ of the set $B_m$ is proportional to $e(|x|)$, where $e(x) \sim c \cdot x$ for a constant $c$, with $0 < c < \frac{1}{2}$.*

*Proof.* To prove that $e(x)$ grows at most linearly, we use the substitution $\tilde{e}(x) = \frac{e(x)}{x}$ for $x > 0$, which transforms the recursion (5.6) into

$$\tilde{e}(x-1)(x-1)r(x-1) + \tilde{e}(x+1)(x+1)r(x+1) - 2\tilde{e}(x)x\big(r(x)+1\big) = 0. \tag{5.11}$$

The sequence $\tilde{e}(x)$ converges if and only if $e(x)$ grows at most linearly. Since $e(x)$ is positive by construction, it suffices to prove that $\tilde{e}(x)$ is decreasing.

Consider the function $r'(x) = \frac{(x+1)^2}{3} - \frac{1}{3}$. We have three cases

$$r(x) = \begin{cases} r'(x), & x \equiv 0 \mod 3 \\ r'(x), & x \equiv 1 \mod 3 \\ r'(x) + \frac{1}{3}, & x \equiv 2 \mod 3 \end{cases}$$

We prove the monotonicity of $\tilde{e}(x)$ by induction. Assuming $\tilde{e}(x) < \tilde{e}(x-1)$ for $x \equiv 0 \mod 3$ we show that $\tilde{e}(x+3) < \tilde{e}(x+2) < \tilde{e}(x+1) < \tilde{e}(x)$.

The basis of the induction follows by calculating the first few elements of the sequence.

*Case (i):* Assume $x \equiv 0 \mod 3$ and $\tilde{e}(x) < \tilde{e}(x-1)$. In this case the recursion (5.11) can be written as

$$\tilde{e}(x+1)(x+1)r'(x+1) = 2\tilde{e}(x)x\big(r'(x)+1\big) - \tilde{e}(x-1)(x-1)\left(r'(x-1)+\tfrac{1}{3}\right)$$

Using the induction hypothesis we get

$$\tilde{e}(x+1) < \tilde{f}_0(x) \cdot \tilde{e}(x), \tag{5.12}$$

with

$$\tilde{f}_0(x) = \frac{2x\big(r'(x)+1\big) - (x-1)\big(r'(x-1)+\frac{1}{3}\big)}{(x+1)r'(x+1)}.$$

Using the definition of $r'(x)$, this simplifies to

$$\tilde{f}_0(x) = \frac{x^2+2x}{x^2+2x+1} < 1,$$

which proves $\tilde{e}(x+1) < \tilde{e}(x)$.

*Case (ii):* Assume $x \equiv 1 \mod 3$ and case (i) for $x-1$. Similarly to the previous case rewrite (5.11) as

$$\tilde{e}(x+1)(x+1)\big(r'(x+1)+\tfrac{1}{3}\big) = 2\tilde{e}(x)x\big(r'(x)+1\big) - \tilde{e}(x-1)(x-1)r'(x-1),$$

which, by (5.12) gives

$$\tilde{e}(x+1) < \tilde{f}_1(x) \cdot \tilde{e}(x), \tag{5.13}$$

for

$$\begin{aligned}\tilde{f}_1(x) &= \frac{2x\big(r'(x)+1\big) - \tilde{f}_0(x-1)^{-1}(x-1)r'(x-1)}{(x+1)\big(r'(x+1)+\frac{1}{3}\big)} \\ &= \frac{x^2+3x}{x^2+3x+2} < 1.\end{aligned}$$

*Case (iii):* Finally, assuming $x \equiv 2 \mod 3$ and case (ii) for $x-1$, we get the recursion

$$\tilde{e}(x+1)(x+1)r'(x+1) = 2\tilde{e}(x)x\big(r'(x)+\tfrac{4}{3}\big) - \tilde{e}(x-1)(x-1)r'(x-1).$$

Applying (5.13), gives the inequality

$$\tilde{e}(x+1) < \tilde{f}_2(x) \cdot \tilde{e}(x), \tag{5.14}$$

for the function

$$\begin{aligned}\tilde{f}_2(x) &= \frac{2x\big(r'(x)+\frac{4}{3}\big)-\tilde{f}_1(x-1)^{-1}(x-1)r'(x-1)}{(x+1)r'(x+1)}\\ &= \frac{x^4+7x^3+17x^2+17x}{x^4+7x^3+17x^2+17x+6}<1.\end{aligned}$$

This shows that $\tilde{e}(x)$ is decreasing and therefore convergent, which also means that $e(x)$, the number of particles that stop at vertex $(x, r(x))$ of the boundary of $\mathcal{B}_m$, is asymptotically at most $\sim c\cdot x$. That $c<\frac{1}{2}$ follows since $\tilde{e}(x)<\frac{1}{2}$ for all $x\geq 20$.

To show that $c>0$, that is, $e(x)$ is at least growing linearly, we can use singularity analysis of linear differential equations. For this we split the $e(x)$ into three sequences modulo 3, i.e., for $k\in\mathbb{N}$ write

$$\begin{aligned}e_0(k) &= e(3k)\\ e_1(k) &= e(3k+1)\\ e_2(k) &= e(3k+2).\end{aligned}$$

As in the previous part of the proof, we can rewrite (5.11) for each congruence class of $x$ mod 3 in terms of $k$. This leads to a system of linear recursions with polynomial coefficients for the sequences $e_i(k)$

$$\begin{aligned}(3k^2+2k)e_0(k)+(3k^2-2k)e_1(k-1)-(6k^2+2)e_2(k-1)&=0\\ (6k^2+4k+2)e_0(k)-(3k^2+4k+1)e_1(k)-3k^2e_2(k-1)&=0\\ (3k^2+2k)e_0(k)-(6k^2+8k+4)e_1(k)+(3k^2+6k+3)e_2(k)&=0.\end{aligned}$$

This can be written in matrix form as

$$A_k\cdot\vec{e}(k-1)=B_k\cdot\vec{e}(k). \tag{5.15}$$

Here $\vec{e}(k)$ denotes the vector $\big(e_0(k),e_1(k),e_2(k)\big)^t$, and the matrices $A_k$ and $B_k$ are given as

$$A_k=\begin{pmatrix}0 & 3k^2-2k & -6k^2-2\\ 0 & 0 & 3k^2\\ 0 & 0 & 0\end{pmatrix}$$

and

$$B_k=\begin{pmatrix}-3k^2-2k & 0 & 0\\ 6k^2+4k+2 & -3k^2-4k-1 & 0\\ 3k^2+2k & -6k^2-8k-4 & 3k^2+6k+3\end{pmatrix}.$$

The initial values are given by $\vec{e}(0)=\big(1,1,\frac{4}{3}\big)^t$.

Denote by $E_i(z) = \sum_{k\geq 0} e_i(k)z^k$ the generating function of $e_i(k)$. The recursion (5.15) can be transformed into a system of linear differential equations for the generating functions $E_i(k)$, using the identities

$$\begin{aligned}\sum_{k\geq 0} ke_i(k)z^k &= z\frac{\partial}{\partial z}E_i(z),\\ \sum_{k\geq 0} k^2e_i(k)z^k &= z^2\frac{\partial^2}{\partial z^2}E_i(z) + z\frac{\partial}{\partial z}E_i(z).\end{aligned}$$

This leads to the following differential equation

$$C \cdot \vec{E}(z) = b, \tag{5.16}$$

where $\vec{E}(z) = \big(E_0(z), E_1(z), E_2(z)\big)^t$, and $C$ is a matrix of linear differential operators given as

$$C = \begin{pmatrix} 5\frac{\partial}{\partial z} + 3z\frac{\partial^2}{\partial z^2} & 1 + 7z\frac{\partial}{\partial z} + 3z^2\frac{\partial^2}{\partial z^2} & -8 - 18z\frac{\partial}{\partial z} - 6z^2\frac{\partial^2}{\partial z^2} \\ -2 - 10z\frac{\partial}{\partial z} - 6z^2\frac{\partial^2}{\partial z^2} & 1 + 7z\frac{\partial}{\partial z} + 3z^2\frac{\partial^2}{\partial z^2} & 3z + 9z^2\frac{\partial}{\partial z} + 3z^3\frac{\partial^2}{\partial z^2} \\ 5z\frac{\partial}{\partial z} + 3z^2\frac{\partial^2}{\partial z^2} & -4 - 14z\frac{\partial}{\partial z} - 6z^2\frac{\partial^2}{\partial z^2} & 3 + 9z\frac{\partial}{\partial z} + 3z^2\frac{\partial^2}{\partial z^2} \end{pmatrix}.$$

The righthand side vector $b$ is equal to

$$b = \begin{pmatrix} 0 \\ e_1(0) - 2e_0(0) \\ 0 \end{pmatrix}.$$

To solve (5.16) asymptotically, we consider $C$ as a matrix with entries in the Weyl algebra, that is, the noncommutative ring of linear differential operators with polynomial coefficients, see [Lam91]. We can perform a division-free Gauss elimination over this ring to transform $C$ into row echelon form, which gives a single differential equation only involving $E_2(z)$. The actual computations were performed using the computer algebra system `FriCAS`[1]. The

[1] `http://fricas.sourceforge.net`

result is a differential equation of order 7 for $E_2(z)$.

$$\begin{aligned}
&\frac{81}{8}(z+2)(z-1)^5 z^6 \frac{\partial^7}{\partial z^7}\\
&+\frac{1269}{4}(z-1)^4 z^5\left(z^2+z-\frac{76}{47}\right)\frac{\partial^6}{\partial z^6}\\
&+\frac{27531}{8}(z-1)^3 z^4\left(z^3-\frac{24}{437}z^2-\frac{7149}{3059}z+\frac{3826}{3059}\right)\frac{\partial^5}{\partial z^5}\\
&+\frac{127725}{8}\cdot(z-1)^2\cdot z^3\cdot\left(z^4-\frac{50039}{42575}z^3-\frac{82401}{42575}z^2+\frac{132307}{42575}z-\frac{38554}{42575}\right)\frac{\partial^4}{\partial z^4}\\
&+31785(z-1)z^2\left(z^5-\frac{100697}{42380}z^4-\frac{1164}{10595}z^3+\frac{36215}{8476}z^2-\frac{5651}{1630}z+\frac{6234}{10595}\right)\frac{\partial^3}{\partial z^3}\\
&+23970z\cdot\left(z^6-\frac{117579}{31960}z^5+\frac{114057}{31960}z^4+\frac{15053}{6392}z^3-\frac{208329}{31960}z^2+\frac{59229}{15980}z-\frac{1243}{3995}\right)\frac{\partial^2}{\partial z^2}\\
&+4935\left(z^6-\frac{1354}{329}z^5+\frac{1843}{329}z^4-\frac{4479}{3290}z^3-\frac{12209}{3290}z^2+\frac{1466}{329}z-\frac{32}{329}\right)\frac{\partial}{\partial z}\\
&+105\left(z^5-\frac{494}{105}z^4+\frac{881}{105}z^3-\frac{201}{70}z^2+\frac{4411}{210}z+\frac{1006}{105}\right)=0
\end{aligned} \tag{5.17}$$

Using singularity analysis for linear differential equations, as described in the previous Section, we can derive asympotics of $e_2(k)$. The coefficient of $\frac{\partial^7}{\partial z^7}$ is given by

$$\frac{81}{8}(z+2)(z-1)^5 z^6,$$

hence the dominant non-zero singularity $\xi$ is equal to 1. Since all coefficients in (5.17) are given in factorized from, it is immediate that $\xi$ is a regular singularity. Calculating the indicial polynomial for the singularity $\xi$ gives

$$I_\xi(\theta) = \theta^7 - 17\theta^6 + 99\theta^5 - 187\theta^4 - 220\theta^3 + 1044\theta^2 - 720\theta.$$

The roots of $I_\xi(\theta)$ are $-2, 0, 1, 3, 4, 5$ and 6, hence by Remark 5.1.4 we might have logarithmic terms in the asympotics. By (5.10) the asymptotics of $e_2(k)$ is given by

$$e_2(k) \sim c \cdot \xi^{-k} k^\beta \log^l k,$$

where $l$ is an integer and $b$ is a biggest solution of the equation $I(-\beta-1) = 0$. In our case $\beta = 1$, and we have

$$e_2(k) \sim c \cdot k \log^l k.$$

While it is not known how to calculate the constant $l$ in the general case, from the first part of the proof we already know that $e_2(k)$ grows at most linearly, hence $l = 0$. □

# Chapter 6

# Internal Diffusion Limited Aggregation on the Comb

This chapter is devoted to the proof of Theorem 2.2.7. The proof uses ideas of Lawler, Bramson and Griffeath [LBG92] and of Levine and Peres [LP10].

For each $n \geq 0$, denote by $\mathcal{B}_n$ the subset of the comb given by

$$\mathcal{B}_n = \left\{ (x, y) \in \mathcal{C}_2 : \frac{|x|}{k} + \left( \frac{|y|}{l} \right)^{1/2} \leq n^{1/3} \right\},$$

where the constants $k$ and $l$ are given by $k = \left(\frac{3}{2}\right)^{2/3}$ and $l = \frac{1}{2}\left(\frac{3}{2}\right)^{1/3}$.

Let $\{X_t^i\}_{i \in \mathbb{N}}$ be a sequence of independent simple random walks on the comb, with common starting point $X_0^i = o$, and $A_n$ be the IDLA-cluster of $n$ particles, as defined in Definition 2.2.5. Following [LBG92], let us introduce some stopping times and random variables. Define

$$\sigma^i = \min\{t \geq 0 : X_t^i \notin A_{i-1}\},$$

which represents the time it takes the $i$-th particle to leave the occupied cluster $A_{i-1}$, and, for $z \in \mathcal{B}_n$,

$$\tau_z^i = \min\{t \geq 0 : X_t^i = z\}$$

denotes the time of the first visit of the $i$-th walk to vertex $z$. The first exit time of the set $\mathcal{B}_n$ is given by

$$\tau_n^i = \min\{t \geq 0 : X_t^i \notin \mathcal{B}_n\}.$$

We want to find an upper bound for the probability that a vertex $z \in \mathcal{B}_n$ does not belong to the IDLA cluster $A_n$, which can be written in terms of the stopping times defined above as

$$\mathbb{P}[z \notin A_n] = \mathbb{P}\left[\bigcap_{i\le n} \sigma^i < \tau_z^i\right].$$

By the Borel-Cantelli Lemma

$$\sum_{n\ge n_0} \sum_{z\in\mathcal{B}_{n(1-\varepsilon)}} \mathbb{P}[z \notin A_n] < \infty, \tag{6.1}$$

is a sufficient condition for proving Theorem 2.2.7. Now fix $n$ and $z \in \mathcal{B}_n$ and consider the random variables

$$N = \sum_{i=1}^{n} \mathbf{1}_{\{\tau_z^i<\sigma^i\}} \quad \text{and} \quad M = \sum_{i=1}^{n} \mathbf{1}_{\{\tau_z^i<\tau_n^i\}},$$

where $N$ represents the number of particles that visit $z$ before leaving the cluster, and $M$ counts the number of particles that visit $z$ before leaving $\mathcal{B}_n$. Let

$$L = \sum_{i=1}^{n} \mathbf{1}_{\{\sigma^i\le\tau_z^i<\tau_n^i\}}$$

be the number of particles that visit $z$ after leaving the cluster $A_i$, but before leaving $\mathcal{B}_n$.

Remark that if $L < M$, then $z$ belongs to the occupied cluster $A_n$. Moreover $N \ge M - L$. Therefore, in order to estimate the probability that $z$ does not belong to the cluster $A_n$, we just need to bound the probability that $M = L$. For any fixed number $a$

$$\begin{aligned}\mathbb{P}[z \notin A_n] &= \mathbb{P}[N = 0] \le \mathbb{P}[M - L = 0]\\ &\le \mathbb{P}[M \le a \text{ or } L \ge a] \le \mathbb{P}[M \le a] + \mathbb{P}[L \ge a].\end{aligned} \tag{6.2}$$

We will show that for a suitable choice of $a$, the probabilities $\mathbb{P}[M \le a]$ and $\mathbb{P}[L \ge a]$ can be made small enough such that the series in (6.1) converges, which implies the inner bound of $A_n$ in Theorem 2.2.7.

The derivation of a suitable value of $a$ needs to be done in a different way, not as in the case of Euclidean lattices, studied by Lawler, Bramson and Griffeath in [LBG92], or Levine and Peres [LP10]. They used classical Green kernel asymptotics available on $\mathbb{Z}^d$, for $d \ge 3$ (see Lawler [Law91]), and asympotics for the recurrent potential kernel (see Spitzer [Spi76]) in the case of $d = 2$. Such estimates are not available on $\mathcal{C}_2$, but one can use the odometer function for the divisible sandpile which we derived in Chapter 3 in order to get enough information about the simple random walk on the comb, and to give good approximations for the Green function stopped at $\mathcal{B}_n$, at least for some special cases.

Recall that $G_n(y,z) = G_{\mathcal{B}_n}(y,z)$ is the Green function stopped on the set $\mathcal{B}_n$. The random variable $M$ is a sum of i.i.d. indicator variables, with

$$\mathbb{E}[M] = n\mathbb{P}_o[\tau_z < \tau_n] = n\frac{G_n(o,z)}{G_n(z,z)}. \tag{6.3}$$

The random variable $L$ is not as easy to estimate directly because it is a sum of indicator random variables which are neither independent nor indentically distributed. Following, [LBG92], we can bound $L$ by a sum of independent indicator random variables as follows. Only those indices $i$ with $X^i_{\sigma^i} \in \mathcal{B}_n$ contribute to $L$ and, for each $y \in \mathcal{B}_n$, there is at most one index $i$ with $X^i_{\sigma^i} = y$. The corresponding post-$\tau_y$ random walks are independent. In order to avoid dependencies in $L$, enlarge the index set to all of $\mathcal{B}_n$ and define

$$\tilde{L} = \sum_{y\in\mathcal{B}_n} \mathbf{1}^y_{\{\tau_z<\tau_n\}},$$

where the indicators $\mathbf{1}^y$ correspond to independent post $\tau_y$ random walks starting at $y$. Then $L \leq \tilde{L}$, and the expectation of $\tilde{L}$ is given by

$$\mathbb{E}[\tilde{L}] = \sum_{y\in\mathcal{B}_n} \mathbb{P}_y[\tau_z < \tau_n] = \frac{1}{G_n(z,z)} \sum_{y\in\mathcal{B}_n} G_n(y,z). \tag{6.4}$$

Now (6.2) can be rewritten as

$$\mathbb{P}[z \notin A_n] \leq \mathbb{P}[M \leq a] + \mathbb{P}[\tilde{L} \geq a]. \tag{6.5}$$

Next we relate the random variables $\tilde{L}$ and $M$ with the odometer function of the divisible sandpile. Consider the function $f_n : \mathcal{B}_n \to \mathbb{R}$, given as

$$f_n(z) = \frac{G_n(z,z)}{d(z)}\mathbb{E}[M - \tilde{L}]. \tag{6.6}$$

Using (6.3) and (6.4), we have

$$\begin{aligned} f_n(z) &= \frac{G_n(z,z)}{d(z)}\Big(n\mathbb{P}[\tau_z < \tau_n] - \sum_{y\in\mathcal{B}_n} \mathbb{P}_y[\tau_z < \tau_n]\Big) \\ &= \frac{1}{d(z)} \sum_{y\in\mathcal{B}_n} \big(n \cdot \delta_o(y) - 1\big) G_n(y,z), \end{aligned}$$

where $o = (0,0)$ is the origin of $\mathcal{C}_2$. By reversibility of the simple random walk on $\mathcal{C}_2$, we have that

$$\frac{G_n(x,y)}{d(y)} = \frac{G_n(y,x)}{d(x)}$$

and the Laplacian of the function $h_x(y) = \frac{G_n(x,y)}{d(y)}$, for a fixed $x$, is given by

$$\triangle h_x(y) = \triangle \frac{G_n(x,y)}{d(y)} = \begin{cases} -\dfrac{1}{d(x)}, & \text{if } x = y \\ 0, & \text{if } x \neq y \end{cases}.$$

Therefore, by linearity, the Laplacian of $f_n$ is equal to

$$\triangle f_n(z) = \frac{1}{d(z)}\big(1 - n \cdot \delta_o(z)\big), \text{ for } z \in \mathcal{B}_n.$$

Thus $f_n$ solves the following Dirichlet problem

$$\begin{cases} \triangle f_n(z) = \dfrac{1}{d(z)}\big(1 - n \cdot \delta_o(z)\big), & \text{for } z \in \mathcal{B}_n \\ f_n(z) = 0, & \text{for } z \in \partial\mathcal{B}_n \end{cases}$$

with $\partial\mathcal{B}_n = \big\{x \in G : \exists y \sim x \text{ and } y \in \mathcal{B}_n\big\}$.

Recall that the divisible sandpile odometer function $u_n$ with initial mass distribution $\mu_0 = n \cdot \delta_o$ solves the same Dirichlet problem on the set $S_n$ of fully occupied sites (see page 22), which is described in Theorem 2.2.2. Since the solution of a Dirichlet problem is unique, it follows that $f_n = u_n$ on the set $\mathcal{B}_n$, and $u_n$ is approximated (up to an additive constant) by the function $\gamma_n$ defined in (3.15). Since $u_n > 0$, it follows that $f_n(z) > 0$, for all $z \in \mathcal{B}_n$ and $\mathbb{E}[M - \tilde{L}] > 0$, that is, $\mathbb{E}[M] > \mathbb{E}[\tilde{L}]$.

We will use the following large deviations estimate for sums of independent indicators. For a proof, see SPENCER [AS92, Cor. $A$.14].

**Lemma 6.0.2.** *If $N$ is a sum of finitely many independent indicator random variables, then for all $\lambda > 0$,*

$$\mathbb{P}\big[|N - \mathbb{E}[N]| > \lambda\mathbb{E}[N]\big] < 2e^{-c_\lambda \mathbb{E}[N]},$$

*where $c_\lambda$ is a constant depending only on $\lambda$.*

In order to find an upper bound for the right-hand side of (6.5) we use the previous Lemma and choose $\lambda > 0$ such that

$$(1+\lambda)\mathbb{E}[\tilde{L}] \leq a \leq (1-\lambda)\mathbb{E}[M], \tag{6.7}$$

for some $a$. Such a $\lambda$ has to satisfy the relation

$$0 < \lambda \leq \frac{\mathbb{E}[M - \tilde{L}]}{\mathbb{E}[M + \tilde{L}]}. \tag{6.8}$$

Define a function $g_n$ as

$$g_n(z) = \frac{G_n(z,z)}{d(z)} \mathbb{E}[M + \tilde{L}], \tag{6.9}$$

then from (6.6) and (6.9) we have

$$\frac{f_n(z)}{g_n(z)} = \frac{\mathbb{E}[M - \tilde{L}]}{\mathbb{E}[M + \tilde{L}]}. \tag{6.10}$$

Since we need a $\lambda$ which satisfies (6.8), it suffices to find the maximal subset of $\mathcal{B}_n$ on which the quotient $f_n(z)/g_n(z)$ is bounded away from 0. For this, we prove that the function $g_n$ solves a Dirichlet problem, and we determine explicitely the unique solution.

## 6.1 The function $g_n$

In this section we compute the function $g_n$. By (6.3) and (6.4) the discrete Laplacian of $g_n$ is equal to

$$\triangle g_n(z) = \frac{1}{d(z)}\big( -1 - n \cdot \delta_o(z)\big), \quad \text{for } z \in \mathcal{B}_n, \tag{6.11}$$

and $g_n(z) = 0$ for all $z \in \partial\mathcal{B}_n$. Therefore, $g_n$ is the unique solution of the Dirichlet problem (6.11).

For simplicity, we first shift the set $\mathcal{B}_n$ defined in (2.3) by $kn^{1/3}$ on the $x$-coordinate. This shifted set will be denoted by $\mathcal{B}_n^t$, which is the set of all $(x,y) \in \mathcal{C}_2$ which satisfy the following three relations

$$0 \leq x \leq 2kn^{1/3}$$

and

$$\begin{aligned} y &\leq \frac{x^2}{3}, \quad \text{for } 0 \leq x \leq kn^{1/3}, \\ y &\leq \frac{(2kn^{1/3} - x)^2}{3}, \quad \text{for } kn^{1/3} < x \leq 2kn^{1/3}. \end{aligned}$$

By this translation, we move the center $(0,0)$ of the set $\mathcal{B}_n$ to $(kn^{1/3}, 0)$ and the left and right corners of $\mathcal{B}_n^t$ are $(0,0)$ and $(2kn^{1/3}, 0)$, respectively.

With the shifted set $\mathcal{B}_n^t$ we can associate the function $g_n^t : \mathcal{B}_n^t \to \mathbb{R}$, by

$$g_n^t(x,y) = g_n(x + kn^{1/3}, y), \tag{6.12}$$

which solves the same Dirichlet problem (6.11) on the set $\mathcal{B}_n^t$ with the origin $o$ moved to the point $(kn^{1/3}, 0)$. By symmetry of $g_n$ it is enough to compute $g_n^t(z)$ for $z = (x, y) \in \mathcal{B}_n^t$, with $0 \leq x \leq kn^{1/3}$ and $y \geq 0$.

For $z = (x, y) \in \mathcal{B}_n^t$, with $y \neq 0$, the Laplacian $\triangle g_n^t(z)$ is equal to $-1/2$, hence on each "tooth" of the comb, $g_n^t$ satisfies the linear recursion

$$2g_n^t(x, y) = g_n^t(x + 1, y) + g_n^t(x, y - 1) + 1,$$

which has the general solution

$$g_n^t(x, y) = \frac{1}{2}(y - y^2) + c_1(x) + yc_2(x), \tag{6.13}$$

where $c_1(x)$ and $c_2(x)$ are functions of $x$, which will be determined in the following. For points $(x, 0), (x, 1) \in \mathcal{C}_2$, we have

$$g_n^t(x, 0) = c_1(x) \quad \text{and} \quad g_n^t(x, 1) = c_1(x) + c_2(x). \tag{6.14}$$

From (6.11) we get the following boundary conditions

$$\begin{aligned} g_n^t(0, 0) &= 0, \\ g_n^t(2kn^{1/3}, 0) &= 0 \end{aligned}$$

and for $0 \leq x \leq kn^{1/3}$

$$g_n^t(x, x^2/3) = 0.$$

On the other hand, from equation (6.13), we get

$$g_n^t(x, x^2/3) = \frac{x^2}{6}\Big(1 - \frac{x^2}{3}\Big) + c_1(x) + \frac{x^2}{3}c_2(x) = 0,$$

which implies that the function $c_2(x)$ can be written as

$$c_2(x) = \frac{1}{2}\Big(\frac{x^2}{3} - 1\Big) - \frac{3}{x^2}c_1(x). \tag{6.15}$$

Moreover, on the $x$-axes the Laplace operator of $g_n^t$ satisfies

$$\begin{aligned} \triangle g_n^t(x, 0) &= \frac{1}{4}\big(g_n^t(x - 1, 0) + g_n^t(x + 1, 0) + 2g_n^t(x, 1)\big) - g_n^t(x, 0) \\ &= \begin{cases} -\frac{1}{4}, & \text{if } x \neq kn^{1/3} \\ -\frac{1}{4}(n + 1), & \text{if } x = kn^{1/3}. \end{cases} \end{aligned} \tag{6.16}$$

For $x \neq kn^{1/3}$, that is, in case that the vertex $(x, 0)$ is not the center point of the set $\mathcal{B}_n^t$, we have

$$g_n^t(x + 1, 0) = 4g_n^t(x, 0) - g_n^t(x - 1, 0) - 2g_n^t(x, 1) - 1$$

and using (6.14) we obtain the first relation between the functions $c_1(x)$ and $c_2(x)$, namely

$$c_1(x+1) = 2c_1(x) - c_1(x-1) - 2c_2(x) - 1$$

Plugging $c_2(x)$, as obtained in (6.15), into the previous relation we obtain a new recursion for the function $c_1$,

$$c_1(x+1) = \Big(2 + \frac{6}{x^2}\Big)c_1(x) - c_1(x-1) - \frac{x^2}{3},$$

which can be explicitely solved and the solution is a polynomial of degree 4, given by

$$c_1(x) = -\frac{1}{18}x^4 + bx^3 - \frac{1}{36}x^2, \tag{6.17}$$

where $b$ is a free parameter which can be computed using the other boundary conditions for $g_n^t$. Since $\triangle g_n^t(kn^{1/3}, 0) = -\frac{1}{4}(n+1)$, using equations (6.14), (6.15), (6.16), and (6.17), we obtain

$$b = \frac{5K + 27n}{18(1+3K^2)},$$

where $K = kn^{1/3}$, and the constant $k = \big(\frac{3}{2}\big)^{2/3}$ is the same as in Theorem 2.2.7. Since we are interested in the form of $g_n^t$, for $n$ sufficiently large, we use the expansion for $b$ around $n = \infty$

$$b(n) = \frac{1}{6l}n^{1/3} + O(n^{-1/3}),$$

and we get the following expression for $g_n^t(z)$,

$$g_n^t(x,y) = \Big(\frac{1}{6l}n^{1/3} + O(n^{-1/3})\Big)(x^3 - 3xy) \\ + \frac{1}{36}(3y - 18y^2 - 2x^4 - x^2 + 12x^2y).$$

Finally, we need to shift the coordinate system back, which gives

$$g_n(x,y) = g_n^t\big(kn^{1/3} - |x|, |y|\big).$$

## 6.2 Proof of the inner bound

Recall that for $n$ big enough, we search for subsets of $\mathcal{B}_n$ on which the inequality

$$0 < \lambda \leq \frac{\mathbb{E}[M - \tilde{L}]}{\mathbb{E}[M + \tilde{L}]}$$

is satisfied, where $\lambda$ is a constant that is independent of the number of particles $n$.

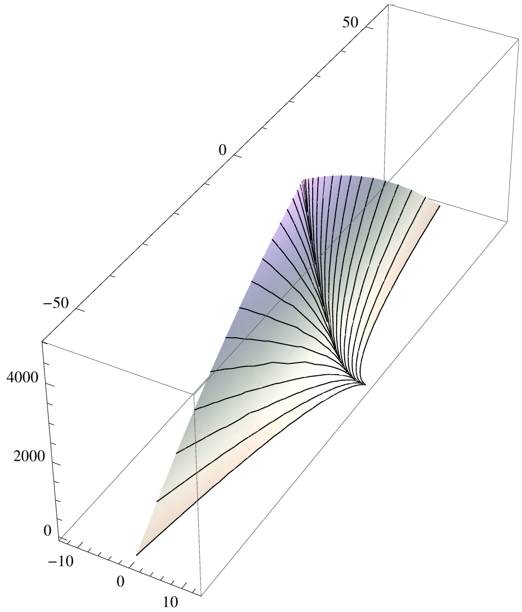



Figure 6.1: Plot of $g_n$ for $n = 1000$, with superimposed $\mathcal{B}_n$ contour lines.

**Lemma 6.2.1.** *For all $\varepsilon > 0$ there exists an $n_\varepsilon$ such that, for all $n \geq n_\varepsilon$ and all $z \in \mathcal{B}_{n(1-\varepsilon)}$,*

$$\frac{\varepsilon}{4} \leq \frac{\mathbb{E}[M - \tilde{L}]}{\mathbb{E}[M + \tilde{L}]}.$$

*Proof.* By (6.10) one needs to study the function $\lambda_n(x, y) = \frac{f_n(x,y)}{g_n(x,y)}$. We have

$$\lambda_n(x, y) = \frac{\left(y - \frac{n_x}{2}\right)^2}{2c_1(x) + \left(2c_2(x) + 1\right)y - y^2},$$

where $c_1(x)$, $c_2(x)$ and $n_x$ are defined in (6.17), (6.15) and (3.12). For every fixed $x \in \mathcal{B}_n$, the function $\lambda_n(x, y)$ is decreasing if $0 \leq y \leq \frac{n_x}{2}$.

From the proof of Theorem 2.2.2 (see (3.17)) we already know that

$$\frac{n_x}{2} = l \left(n^{1/3} - \frac{x}{k}\right)^2 + \mathcal{O}(1). \tag{6.18}$$

For $0 < \varepsilon < 1$ consider the set

$$\mathcal{B}_{n,\varepsilon} = \Bigg\{(x, y) \in \mathcal{C}_2 : |x| \leq (1 - \varepsilon)kn^{1/3} \text{ and } |y| \leq (1 - \varepsilon)l \left(n^{1/3} - \frac{|x|}{k}\right)^2 \Bigg\}.$$

Obviously $\mathcal{B}_{n,\varepsilon} \subset \mathcal{B}_n$ for all $\varepsilon$, hence $\frac{f_n}{g_n}$ is well defined on this set. Furthermore by (6.18), $\frac{f_n}{g_n}$ is also decreasing on $\mathcal{B}_{n,\varepsilon}$ as a function of $y$, for all $\varepsilon > 0$ and $n$ big enough. This means that it is enough to study $\frac{f_n}{g_n}$ at the inner boundary of $\mathcal{B}_{n,\varepsilon}$.

Let $z = (x, y) \in \mathcal{B}_{n,\varepsilon}$ with $|y| = (1-\varepsilon)l\left(n^{1/3} - \frac{|x|}{k}\right)^2$ be such a boundary point. Then one can calculate the limit

$$\lim_{n\to\infty} \frac{f_n(z)}{g_n(z)} = \frac{\varepsilon}{4-\varepsilon} > \frac{\varepsilon}{4}.$$

The lemma follows from the fact that for each $\varepsilon > 0$ one can find an $\varepsilon' > 0$ such that $\mathcal{B}_{n(1-\varepsilon)} \subset \mathcal{B}_{n,\varepsilon'}$ □

With this in hand, we are now ready to prove Theorem 2.2.7.

*Proof.* Recall that we need to show the convergence of the series (6.1). We already established in (6.5) that

$$\mathbb{P}[z \notin A_n] \leq \mathbb{P}[M \leq a] + \mathbb{P}[\tilde{L} \geq a],$$

where $a$ has the property (6.7). By Lemma 6.2.1, set $\lambda = \frac{\varepsilon}{4} > 0$, then by (6.7) we can choose

$$a = (1+\lambda)\mathbb{E}[\tilde{L}] = \Big(1 + \frac{\varepsilon}{4}\Big)\mathbb{E}[\tilde{L}].$$

For this choice of the constant $a$ apply Lemma 6.0.2 to $M$ and $\tilde{L}$. Recall also that $\mathbb{E}[M] > \mathbb{E}[\tilde{L}]$. Then

$$\begin{aligned}
\mathbb{P}[M \leq a] + \mathbb{P}[\tilde{L} \geq a] &= \mathbb{P}\Big[M \leq \Big(1+\frac{\varepsilon}{4}\Big)\mathbb{E}[\tilde{L}]\Big] + \mathbb{P}\Big[\tilde{L} \geq \Big(1+\frac{\varepsilon}{4}\Big)\mathbb{E}[\tilde{L}]\Big] \\
&< 2\exp\big\{-c_\lambda \mathbb{E}[\tilde{L}]\big\} + 2\exp\big\{-c_\lambda \mathbb{E}[\tilde{L}]\big\} \\
&= 4\exp\Big\{-c_\lambda d(z)\frac{g_n(z) - f_n(z)}{2G_n(z,z)}\Big\} \\
&\leq 4\exp\Big\{-c_\lambda \frac{g_n(z) - f_n(z)}{G_n(z,z)}\Big\},
\end{aligned}$$

where $c_\lambda$ is a constant only depending on $\lambda$. Hence, for all $n \geq n_\varepsilon$, we have

$$\sum_{n\geq n_\varepsilon} \sum_{z\in\mathcal{B}_{n(1-\varepsilon)}} \mathbb{P}[z \notin A_n] \leq 4 \sum_{n\geq n_\varepsilon} \sum_{z\in\mathcal{B}_{n(1-\varepsilon)}} \exp\Big\{-c_\lambda \frac{g_n(z) - f_n(z)}{G_n(z,z)}\Big\}, \quad (6.19)$$

and we have to prove that the series on the right hand-side converges. In order to find an upper bound for the stopped Green function $G_n(z,z)$ upon exiting $\mathcal{B}_n$, with $z = (x, y)$, note that

$$|y| \leq b_n(x) := l\Big(n^{1/3} - \frac{|x|}{k}\Big)^2.$$

An upper bound for $G_n(z,z)$ is two times the stopped Green function $G_A(y,y)$ for the simple random walk on the integer line upon leaving the finite symmetric interval $A = \big[-b_n(x), b_n(x)\big]$. Using Proposition 1.6.3 and Theorem 1.6.4 from Lawler [Law91], this can be estimated from above by

$$G_A(y,y) \leq \frac{l^2\big(n^{1/3} - \frac{|x|}{k}\big)^4 - y^2}{l\big(n^{1/3} - \frac{|x|}{k}\big)^2} \leq l\Big(n^{1/3} - \frac{|x|}{k}\Big)^2. \tag{6.20}$$

For every $\varepsilon > 0$, the function $g_n(z) - f_n(z)$ is again decreasing on every non-crossing path which starts at $o$ and which is restricted to $\mathcal{B}_{n(1-\varepsilon)}$. Hence, it attains its minium on the inner boundary of $\mathcal{B}_{n(1-\varepsilon)}$. Taking limits, we get for every sequence $z_n = (x, y_n)$ with $x$ fixed and $z_n \in \partial_I \mathcal{B}_{n(1-\varepsilon)}$

$$\lim_{n\to\infty} \frac{g_n(z_n) - f_n(z_n)}{n^{4/3}} = \frac{k}{4}(2-\varepsilon)\varepsilon,$$

and for the sequence $z'_n = (x_n, 0)$ with $x_n = kn^{1/3}(1-\varepsilon)^{1/3}$

$$\lim_{n\to\infty} \frac{g_n(z'_n) - f_n(z'_n)}{n^{4/3}} = \frac{k}{4}\big(3 - 2\varepsilon - (\varepsilon - 3)(1-\varepsilon)^{1/3}\big).$$

Hence for all $\varepsilon > 0$ and $n$ big enough

$$\min_{z\in\mathcal{B}_{n(1-\varepsilon)}} \big(g_n(z) - f_n(z)\big) \geq C_\varepsilon \cdot n^{4/3},$$

for a constant $C_\varepsilon$ which depends only on $\varepsilon$. Since, by (6.20) the stopped Green function $G_A(z,z)$ is of order $\mathcal{O}(n^{2/3})$, this implies

$$\min_{z\in\mathcal{B}_{n(1-\varepsilon)}} \frac{g_n(z) - f_n(z)}{G_n(z,z)} \geq C'_\varepsilon \cdot n^{2/3}.$$

Hence with (6.19) we get

$$\begin{aligned}
\sum_{n\geq n_\varepsilon} \sum_{z\in\mathcal{B}_{n(1-\varepsilon)}} \mathbb{P}[z \notin A_n] &\leq 4 \sum_{n\geq n_\varepsilon} \sum_{z\in\mathcal{B}_{n(1-\varepsilon)}} \exp\left\{-c_\lambda \min_{z\in\mathcal{B}_{n(1-\varepsilon)}} \frac{g_n(z) - f_n(z)}{G_n(z,z)}\right\} \\
&\leq 4 \sum_{n\geq n_\varepsilon} n \exp\{-c_\lambda C'_\varepsilon n^{2/3}\} < \infty,
\end{aligned}$$

which proves the inner bound

$$\mathbb{P}\big[\mathcal{B}_{n(1-\varepsilon)} \subset A_n, \text{ for all } n \geq n_\varepsilon\big] = 1.$$

$\square$

## Chapter 7

# IDLA on Nonamenable Graphs

In this chapter we are again interested in the shape of IDLA-clusters. We will show that, if the underlying random walk is highly transient and if it satisfies some additional regularity conditions, the limiting shape of the IDLA-clusters does not depent on the local geometry of the graph $G$, but only on the structure of the Green function. This is an extension of a result of BLACHÈRE AND BROFFERIO [BB07].

## 7.1 Random Walks on Nonamenable Graphs

First, we will define what we understand by a random walk on a nonamenable graph. For this we need some new terminology in addition to the notions already introduced in Section 2.1. We follow again the book of WOESS [Woe00].

The *spectral radius* is defined as

$$\rho(P) = \limsup_{t \to \infty} p^{(t)}(x, y)$$

and is independent of $x, y \in G$, because of irreducibility. It is known that

$$p^{(t)}(x, x) \le \rho(P)^t, \text{ for all } x \in G. \tag{7.1}$$

For a random walk $X_t$, let $\tau_y$ be the first hitting time of the point $y$:

$$\tau_y = \inf \{t \ge 0 : X_t = y\},$$

with the convention $\tau_y = \infty$, if the random walk never visits $y$. The hitting probability of $y$, for a random walk starting in $x$, is denoted by

$$F(x, y) = \mathbb{P}^x [\tau_y < \infty]. \tag{7.2}$$

A simple calculation shows

$$G(x,y) = F(x,y)G(y,y).$$

**Definition 7.1.1.** *A random walk is called* uniformly irreducible*, if there are constants* $\varepsilon_0 > 0$ *and* $T < \infty$*, such that*

$$x \sim y \text{ implies } p^{(t)}(x,y) \geq \varepsilon_0, \text{ for some } t \leq T.$$

**Definition 7.1.2.** *A Markov chain* $(G,P)$ *is called* reversible*, if there exists a measure* $m: G \to (0,\infty)$*, such that*

$$m(x)p(x,y) = m(y)p(y,x), \text{ for all } x,y \in G.$$

*If* $m$ *is bounded, i.e., there exists a* $C \in (0,\infty)$*, such that*

$$C^{-1} \leq m(x) \leq C, \text{ for all } x \in X, \tag{7.3}$$

*then the Markov chain is called* strongly reversible.

**Remark 7.1.3.** *The simple random walk is strongly reversible if and only if the vertex degree of* $G$ *is bounded. In that case it is also uniformly irreducible.*

The following inequality from [Woe00, Lemma 8.1] gives us a generalization of (7.1) for arbitrary $t$-step transition probabilities in the case of strongly reversible Markov chains.

**Lemma 7.1.4.** *If* $(G,P)$ *is strongly reversible then* $p^{(t)}(x,y) \leq C\rho(P)^t$*, with* $C$ *as in Definition 7.1.2.*

The next two propositions give a characterization of uniformly irreducible random walks in terms of the the hitting probability $F$, defined in (7.2).

**Proposition 7.1.5.** *If* $(G,P)$ *is uniformly irreducible, then there exists* $\varepsilon_0 > 0$ *such that, for all neighbours* $x \sim y$

$$F(x,y) \geq \varepsilon_0.$$

*Proof.* Recall from Definition 7.1.1, that there exist $\varepsilon_0 > 0$ and $t \leq T$ such that

$$\varepsilon_0 \leq p^{(t)}(x,y) = \mathbb{P}^x[X_t = y] \leq \mathbb{P}^x[\tau_y < \infty] = F(x,y).$$

□

The converse is not true in general and it is easy to construct counterexamples. Using some additional assumptions we can show the following result.

**Proposition 7.1.6.** *Let $(G, P)$ be a strongly reversible random walk with $\rho(P) < 1$, and such that there exists a constant $\varepsilon_0 > 0$, such that $F(x, y) \geq \varepsilon_0$, for all neighbours $x \sim y \in G$. Then the random walk is uniformly irreducible.*

*Proof.* Suppose that the random walk is not uniformly irreducible. This means that for all $\delta > 0$ and for all $T \in \mathbb{N}$, there exist neighbouring points $x_{\delta,T} \sim y_{\delta,T}$ such that, for all $t \leq T :\ p^{(t)}(x_{\delta,T}, y_{\delta,T}) < \delta$.

So for every $T \in \mathbb{N}$, we can construct two sequences $\{x_{i,T}\}_{i\in\mathbb{N}}$ and $\{y_{i,T}\}_{i\in\mathbb{N}}$, with $x_{i,T} \sim y_{i,T}$ and

$$p^{(t)}(x_{i,T}, y_{i,T}) < \frac{1}{i}, \text{ for all } t \leq T. \tag{7.4}$$

For $t = 0$ this implies $x_{i,T} \neq y_{i,T}$.

Lemma 7.1.4 gives a second bound for the $t$-step transition probabilities

$$p^{(t)}(x_{i,T}, y_{i,T}) \leq C\rho(P)^t. \tag{7.5}$$

Define $T_i = \left\lfloor -\frac{\ln(i\cdot C)}{\ln \rho(P)} \right\rfloor$ and two sequences $\{x_i\}_{i\in\mathbb{N}}$ and $\{y_i\}_{i\in\mathbb{N}}$ by

$$x_i = x_{i,T_i} \quad \text{and} \quad y_i = y_{i,T_i}.$$

Using (7.4) and (7.5) we have the following upper bound of the transition probabilities

$$p^{(t)}(x_i, y_i) \leq \begin{cases} \frac{1}{i}, & \text{for } t \leq T_i \\ C\rho(P)^t, & \text{for } t > T_i. \end{cases} \tag{7.6}$$

Now we look at the Green function

$$G(x_i, y_i) = F(x_i, y_i)G(y_i, y_i) \geq F(x_i, y_i) \geq \varepsilon_0. \tag{7.7}$$

But, on the other hand, using (7.6)

$$\begin{aligned} G(x_i, y_i) = \sum_{t=1}^{\infty} p^{(t)}(x_i, y_i) &\leq \sum_{t=1}^{T_i} \frac{1}{i} + C \sum_{t=T_i+1}^{\infty} \rho(P)^t \\ &\leq -\frac{\ln(i \cdot C)}{i \cdot \ln \rho(P)} + \frac{C}{1-\rho(P)} \rho(P)^{T_i+1}. \end{aligned}$$

This goes to 0 for $i \to \infty$, because $T_i$ goes to infinity and $\rho(P) < 1$, and gives a contradiction to the lower bound of the Green function in equation (7.7). □

For a subset $K \subset X$ set $\partial_E(X) = \{(x,y) \in G^2 : \, x \sim y, \, x \in K, \, y \notin K\}$. The *edge-isoperimetric constant* of $G$ is defined as

$$\iota_E(X) = \inf \left\{ \frac{|\partial_E K|}{|K|} : \, K \subset G \text{ finite}, \, K \neq \emptyset \right\}.$$

A Graph $G$ is called *amenable* if $\iota_E(x) = 0$ and *non-amenable* if $\iota_E(x) > 0$. The following relation between amenability and the spectral radius of simple random walk is well known [Dod84, DK86]. See [Woe00, Theorem 10.3] for the proof of a more general version of this theorem.

**Theorem 7.1.7.** *Let $P$ be the transition operator of the simple random walk on $G$. Then $G$ is non-amenable if and only if $\rho(P) < 1$.*

The following theorem (see [Woe00, Theorem 10.3]) will be needed later. For a reversible Markov chain we define the real Hilbert space $\ell^2(G,m)$ with inner product

$$\langle f, g \rangle = \sum_{x \in G} f(x) g(x) m(x).$$

**Theorem 7.1.8.** *The following statements are equivalent for reversible Markov chains $(G,P)$.*

(a) *The spectral radius $\rho(P)$ is strictly smaller than 1.*

(b) *The Green function defines a bounded linear operator $\mathsf{G}$ on $\ell^2(G,m)$ by*

$$\mathsf{G} f(x) = \sum_{y \in G} G(x,y) f(y).$$

From now on we only consider Markov chains $(G,P)$ that are uniformly irreducible, strongly reversible and have $\rho(P) < 1$. See Nagnibeda and Woess [NW02] for a class of graphs which have these properties. In this setting we can follow [BB07] and define the "*hitting distance*". For all $x, y \in G$ let

$$\mathsf{d}_H(x,y) = -\ln F(x,y).$$

If $G$ is the Cayley graph of a finitely generated group, and $P$ a symmetric left invariant random walk on $G$ then $\mathsf{d}_H$ is a left invariant metric on $G$, see [BB07]. For arbitrary graphs this does not hold anymore, but we can still use $\mathsf{d}_H$ to define balls of radius $Kn$, where $K = -\ln \varepsilon_0$ (with $\varepsilon_0$ as in Definition 7.1.1)

$$\mathcal{B}_n(x) = \{z \in G : \, \mathsf{d}_H(x,z) \leq Kn\},$$

and its boundaries as

$$\partial\mathcal{B}_n(x) = \{y \notin \mathcal{B}_n(x) : \exists z \sim y \wedge z \in \mathcal{B}_n(x)\}.$$

The constant $K$ is needed in the definition of the balls to ensure that the balls with radius $n$ and $n+1$ are properly nested, i.e.,

$$\partial\mathcal{B}_n(x) \subseteq \mathcal{B}_{n+1}(x). \tag{7.8}$$

Indeed, let $y \in \partial\mathcal{B}_n(x)$ and $z \in \mathcal{B}_n(x)$ such that $y \sim z$. Then

$$\begin{aligned} \mathsf{d}_H(x,y) &= -\ln F(x,y) \le -\ln F(x,z)F(z,y) = \mathsf{d}(x,z) - \ln F(z,y) \\ &\le Kn - \ln \varepsilon_0 = K(n+1). \end{aligned}$$

This implies that $y \in \mathcal{B}_{n+1}(x)$.

**Proposition 7.1.9.** *The balls $\mathcal{B}_n(x)$ are finite.*

*Proof.* It suffices to show that

$$\lim_{n\to\infty} F(x,y_n) = 0.$$

for all sequences $\{y_n\}_{n\in\mathbb{N}}$ which only contain distinct elements. The functions $\{e_x\}_{x\in G}$ with $e_x(z) = \delta_x(z)m(x)^{-1/2}$ form an orthonormal basis of $\ell^2(G,m)$. By Theorem 7.1.8 the Green function defines a bounded linear operator $\mathsf{G}$, hence its adjoint $\mathsf{G}^*$ is also a bounded linear operator. By Bessel's inequality

$$\sum_{n\ge 0} |\langle e_{y_n}, \mathsf{G}^* e_x\rangle|^2 \le \|\mathsf{G}^* e_x\|^2,$$

for all $x \in G$. Therefore $\langle \mathsf{G}e_{y_n}, e_x\rangle = \langle e_{y_n}, \mathsf{G}^* e_x\rangle \to 0$, for $n \to \infty$. Because of

$$\langle \mathsf{G}e_{y_n}, e_x\rangle = G(x,y_n)\sqrt{\frac{m(x)}{m(y_n)}},$$

and by strong reversibility

$$C^{-1} \le \sqrt{\frac{m(x)}{m(y_n)}} \le C,$$

this is equivalent to $G(x,y_n) \to 0$. Since

$$F(x,y_n) = \frac{G(x,y_n)}{G(y_n,y_n)} \le G(x,y_n),$$

the finiteness of $\mathcal{B}_n(x)$ follows. □

To simplify the notation we write $\mathcal{B}_n = \mathcal{B}_n(o)$ and $\partial\mathcal{B}_n = \partial\mathcal{B}_n(o)$. Denote by $V(n)$ the size of the ball of radius $n$

$$V(n) = |\mathcal{B}_n|.$$

**Proposition 7.1.10.** *For all finite subsets $A \subset G$ the following estimates hold*

$$\sum_{x\in A} F(x,y) \leq \mathcal{J} \cdot c_\rho^{-1} \ln|A|, \tag{7.9}$$

$$\sum_{y\in A} F(x,y) \leq \mathcal{J} \cdot c_\rho^{-1} \ln|A|, \tag{7.10}$$

*with $c_\rho = -\ln\rho(P)$ and some constant $\mathcal{J} > 0$, which does not depend on $A$.*

*Proof.* We have

$$F(x,y) = \frac{G(x,y)}{G(y,y)} \leq G(x,y).$$

Therefore, using the estimate of Lemma 7.1.4 we can write

$$\begin{aligned}\sum_{x\in A} F(x,y) &\leq \sum_{x\in A}\sum_{t=0}^{\infty} p^{(t)}(x,y) = \sum_{t=0}^{\infty}\sum_{x\in A} p^{(t)}(x,y)\\ &\leq \sum_{t\leq c_\rho^{-1}\ln|A|}\ \sum_{x\in A} \frac{m(y)}{m(x)} p^{(t)}(y,x) + \sum_{t> c_\rho^{-1}\ln|A|}\ \sum_{x\in A} C\cdot\rho(P)^t.\end{aligned}$$

For $M = \sup\left\{\frac{m(x)}{m(y)} : x,y\in X\right\}$ we get

$$\begin{aligned}\sum_{x\in A} F(x,y) &\leq M\cdot c_\rho^{-1}\ln|A| + C|A|\int_{c_\rho^{-1}\ln|A|}^{\infty} e^{-c_\rho t}dt\\ &= M\cdot c_\rho^{-1}\ln|A| + C\cdot c_\rho^{-1}\\ &\leq J\cdot c_\rho^{-1}\ln|A|.\end{aligned}$$

The second estimate can be derived in the same way, but without the need to apply reversibility. □

The next proposition gives some estimates of the size of the balls.

**Proposition 7.1.11.** *There exist constants $\mathcal{C}_l, \mathcal{C}_u > 0$, such that*

$$\mathcal{C}_l e^{Kn} \leq V(n) \leq \mathcal{C}_u n e^{Kn}, \text{ for all } n \geq n_0.$$

*Proof.* Since for $x \notin \mathcal{B}_{n-1}$ the distance $\mathsf{d}_H(o,x) > K(n-1)$, it follows that

$$\sum_{x\in\partial\mathcal{B}_{n-1}} F(o,x) \le |\partial\mathcal{B}_{n-1}| \cdot e^{-K(n-1)}. \tag{7.11}$$

Every random walk that leaves the ball $\mathcal{B}_{n-1}$ has to hit at least one point of $\partial\mathcal{B}_{n-1}(n)$. Because the balls are finite, the random walk leaves $\mathcal{B}_{n-1}$ with probability 1, hence

$$\sum_{x\in\partial\mathcal{B}_{n-1}} F(o,x) \ge 1.$$

Then (7.11) and (7.8) imply

$$e^{K(n-1)} \le |\partial\mathcal{B}_{n-1}| \le |\mathcal{B}_n| = V(n).$$

Choosing $\mathcal{C}_l = e^{-K}$, gives the lower bound of the proposition.

To prove the upper bound, consider the following inequality

$$|\mathcal{B}_n| \cdot e^{-Kn} \le \sum_{x\in\mathcal{B}_n} F(o,x) \le \mathcal{J} \cdot c_\rho^{-1} \cdot \ln|\mathcal{B}_n|, \tag{7.12}$$

which follows from the definition of the balls $\mathcal{B}_n$ and Proposition 7.1.10. The finiteness of the balls (Proposition 7.1.9) implies that $|\mathcal{B}_n| = O(e^{K'n})$. Further, (7.12) gives for some constants $\tilde{C}$ and $\mathcal{C}_u$

$$\begin{aligned} |\mathcal{B}_n| &\le \mathcal{J} \cdot c_\rho^{-1} e^{Kn} \ln|\mathcal{B}_n| \\ &\le \mathcal{J} \cdot c_\rho^{-1} e^{Kn} \ln\left(\tilde{C} e^{K'n}\right) \\ &\le \mathcal{J} \cdot c_\rho^{-1} e^{Kn} \left(K'n + \ln\tilde{C}\right) \\ &\le \mathcal{C}_u n e^{Kn}. \end{aligned}$$

□

## 7.2 Internal Diffusion Limited Aggregation

We can now formulate our main result of this chapter, which connects the shape of the IDLA clusters to the balls with respect to the hitting distance.

**Theorem 7.2.1.** *Let the sequence of random subsets $\{A_n\}_n \in \mathbb{N}$ be the IDLA process on a strongly reversible, uniformly irreducible Markov chain with spectral radius strictly smaller than 1. Then for any $\varepsilon > 0$ and all constants $\mathcal{C}_I \ge \frac{1+\varepsilon}{K}$ (where $K$ is the constant used in the definition of the balls $\mathcal{B}_n$) and $\mathcal{C}_O > \sqrt{3}$,*

$$\mathbb{P}\left[\exists n_\varepsilon \text{ s.t. } \forall n \ge n_\varepsilon : \mathcal{B}_{n-\mathcal{C}_I \ln n} \subseteq A_{V(n)} \subseteq \mathcal{B}_{n+\mathcal{C}_O\sqrt{n}}\right] = 1.$$

The proof of Theorem 7.2.1 is very similar to the proof of the equivalent shape theorem for IDLA on groups with exponential growth [BB07, Theorem 3.1].

Recall that $X_t^j$ represents the $j^{\text{th}}$ random walk of the IDLA process. The main tools of the proof are two types of stopping times. As in Chapter 6, denote by $\sigma^j$ the first time at which the $j^{\text{th}}$ random walk adds to the cluster

$$\sigma^j = \inf\left\{t \geq 0 \; : \; X_t^j \notin A_{j-1}\right\}, \tag{7.13}$$

and by $\tau_z^j$ the time at which the $j^{\text{th}}$ random walk first hits a point $z$

$$\tau_z^j = \inf\left\{t \geq 0 \; : \; X_t^j = z\right\}. \tag{7.14}$$

### 7.2.1 The Inner Bound

To prove the inner bound, we first estimate the probability that a point $z \in \mathcal{B}_n$ is part of the IDLA-cluster which contains $V(n)$ particles.

To do this we, define a few random variables

$$N^j(z) = \mathbf{1}_{\{\tau_z^j \leq \sigma^j\}}, \tag{7.15}$$

$$M^j(z) = \mathbf{1}_{\{\tau_z^j < \infty\}}, \tag{7.16}$$

$$L^j(z) = \mathbf{1}_{\{\sigma^j < \tau_z^j < \infty\}}. \tag{7.17}$$

It is easy to see that the relation $N^j(z) = M^j(z) - L^j(z)$ holds. Since $z$ will be fixed, we will just write $N^j = N^j(z)$, and so on, to simplify the notation.

Note that a point $z$ is not part of the IDLA-cluster at a time $T$, if and only if $N^j = 0$, for all $j \leq T$. So, for all $\lambda \geq 0$

$$\begin{aligned}
\mathbb{P}\left[z \notin A_{V(n)}\right] &= \mathbb{P}\left[\sum_{j=1}^{V(n)} N^j = 0\right] \\
&\leq \sum_{k \geq 0} \mathbb{P}\left[\sum_{j=1}^{V(n)} N^j = k\right] \cdot e^{-\lambda k} = \mathbb{E}\left[e^{-\lambda \sum_{j=1}^{V(n)} N^j}\right] \\
&= \mathbb{E}\left[e^{-\lambda \sum_{j=1}^{V(n)} M^j - L^j}\right] = \mathbb{E}\left[e^{-\lambda \sum_{j=1}^{V(n)} M^j} \cdot e^{\lambda \sum_{j=1}^{V(n)} L^j}\right] \\
&\leq \mathbb{E}\left[e^{-2\lambda \sum_{j=1}^{V(n)} M^j}\right]^{\frac{1}{2}} \cdot \mathbb{E}\left[e^{2\lambda \sum_{j=1}^{V(n)} L^j}\right]^{\frac{1}{2}}.
\end{aligned} \tag{7.18}$$

The random variables $M^j$ are all identically distributed and independent, therefore

$$\begin{aligned}
\mathbb{E}\left[e^{-2\lambda\sum_{j=1}^{V(n)} M^j}\right] &= \prod_{j=1}^{V(n)} \mathbb{E}\left[e^{-2\lambda M^j}\right] \\
&= \prod_{j=1}^{V(n)} \left(\mathbb{P}\left[M^j = 0\right] + \mathbb{P}\left[M^j = 1\right] \cdot e^{-2\lambda}\right) \\
&= \prod_{j=1}^{V(n)} \left((1 - F(z)) + F(z)e^{-2\lambda}\right) \\
&= \left(1 - (1 - e^{-2\lambda})F(z)\right)^{V(n)}.
\end{aligned}$$

Applying the inequality $(1-x)^y \le e^{-xy}$ (which holds for $x < 1$) implies:

$$\mathbb{E}\left[e^{-2\lambda\sum_{j=1}^{V(n)} M^j}\right] \le \exp\left(-(1 - e^{-2\lambda})F(z)V(n)\right).$$

To estimate the second part of the right hand side of (7.18), we bound $L^j$ from above by random variables that are independent when conditioned with respect to a suitable $\sigma$-algebra. Let

$$\tilde{\tau}_z^j = \inf\left\{t \ge \sigma^j \ : \ X_t^j = z\right\},$$

then

$$L^j \le \tilde{L}^j = \mathbf{1}_{\{\tilde{\tau}_z^j < \infty\}}.$$

We define a sequence of $\sigma$-algebras:

$$\mathcal{G}_n = \sigma\left(X_{t\wedge\sigma^j}^j\right), \text{ for all } j \le V(n) \text{ and } t \in \mathbb{N}. \tag{7.19}$$

$\mathcal{G}_n$ encodes all information about the random walks $X_t^j$ before they add to the IDLA-cluster $A_{V(n)}$.

The random variables $\tilde{L}^j$ are independent when conditioned with respect to $\mathcal{G}_n$. Hence

$$\begin{aligned}
\mathbb{E}\left[e^{2\lambda\sum_{j=1}^{V(n)} L^j}\right] \le \mathbb{E}\left[e^{2\lambda\sum_{j=1}^{V(n)} \tilde{L}^j}\right] &= \mathbb{E}\left[\mathbb{E}\left[e^{2\lambda\sum_{j=1}^{V(n)} \tilde{L}^j} \mid \mathcal{G}_n\right]\right] \\
&= \mathbb{E}\left[\prod_{j=1}^{V(n)} \mathbb{E}\left[e^{2\lambda \tilde{L}^j} \mid \mathcal{G}_n\right]\right].
\end{aligned}$$

Since the $\tilde{L}^j$ are indicator random variables, this is further

$$
\begin{aligned}
&= \mathbb{E}\left[\prod_{j=1}^{V(n)}\left(1+(e^{2\lambda}-1)\mathbb{E}\left[\tilde{L}^j \mid \mathcal{G}_n\right]\right)\right] \\
&\leq \mathbb{E}\left[\exp\left((e^{2\lambda}-1)\sum_{j=1}^{V(n)}\mathbb{E}\left[\tilde{L}^j \mid \mathcal{G}_n\right]\right)\right].
\end{aligned}
$$

Finally, we need an estimate of the conditional expectation of $\tilde{L}^j$. With Proposition 7.1.10 we get

$$
\begin{aligned}
\sum_{j=1}^{V(n)}\mathbb{E}\left[\tilde{L}^j \mid \mathcal{G}_n\right] &= \sum_{j=1}^{V(n)}\mathbb{E}\left[\mathbf{1}_{\{\tilde{\tau}_z^j<\infty\}} \mid \mathcal{G}_n\right] = \sum_{j=1}^{V(n)}\mathbb{P}_{X^j_{\sigma^j}}\left[\tilde{\tau}_z^j<\infty\right] \\
&= \sum_{y\in A_{V(n)}} F(y,z) \leq \mathcal{J}\cdot c_\rho^{-1}\ln V(n).
\end{aligned}
$$

Denote by $\mathcal{E}$ the event $\left[z\notin A_{V(n)}\right]$. Putting the pieces together, we get the estimate

$$
\begin{aligned}
\mathbb{P}[\mathcal{E}] &\leq \exp\left(-(1-e^{-2\lambda})F(z)V(n)\right)^{\frac{1}{2}}\cdot\exp\left((e^{2\lambda}-1)\mathcal{J}c_\rho^{-1}\ln V(n)\right)^{\frac{1}{2}} \\
&= \exp\left(\frac{1}{2}\left(-(1-e^{-2\lambda})F(z)V(n)+(e^{2\lambda}-1)\mathcal{J}c_\rho^{-1}\ln V(n)\right)\right).
\end{aligned}
$$

Using the bounds for $V(n)$ from Proposition 7.1.11, we arrive at

$$
\begin{aligned}
\mathbb{P}[\mathcal{E}] &\leq \exp\left(\frac{1}{2}\left(-(1-e^{-2\lambda})F(z)\mathcal{C}_l e^{Kn}+(e^{2\lambda}-1)\mathcal{J}c_\rho^{-1}\ln(\mathcal{C}_u n e^{Kn})\right)\right) \\
&\leq \exp\left(-\mathcal{C}_\lambda F(z)e^{Kn}+\mathcal{C}'_\lambda Kn\right),
\end{aligned}
$$

for some constants $\mathcal{C}_\lambda$, $\mathcal{C}'_\lambda>0$, depending on $\lambda$.

We choose $z\in\mathcal{B}_{n-\mathcal{C}_I\ln n}$ for a positive constant $\mathcal{C}_I$. This means by definition that

$$F(z)\geq e^{K\mathcal{C}_I\ln n-Kn}=n^{\mathcal{C}_I K}\cdot e^{-Kn}.$$

So, for all $\varepsilon>0$, $\mathcal{C}_I\geq\frac{1+\varepsilon}{K}$ and $n\geq n_\varepsilon$:

$$\mathbb{P}\left[z\notin A_{V(n)}\right]\leq\exp\left(-\mathcal{C}_\lambda n^{\mathcal{C}_I K}+\mathcal{C}'_\lambda Kn\right)\leq\exp\left(-\mathcal{C}_\lambda n^{1+\varepsilon}\right).$$

Using the upper bound of Proposition 7.1.11 again, we get

$$
\begin{aligned}
\sum_{z\in\mathcal{B}_{n-\mathcal{C}_I\ln n}}\mathbb{P}\left[z\notin A_{V(n)}\right] &\leq V(n-\mathcal{C}_I\ln n)\exp\left(-\mathcal{C}_\lambda n^{1+\varepsilon}\right) \\
&\leq \mathcal{C}_u(n-\mathcal{C}_I\ln n)\cdot n^{-K\mathcal{C}_I}\cdot e^{Kn}\cdot e^{-\mathcal{C}_\lambda n^{1+\varepsilon}},
\end{aligned}
$$

which is summable. Applying the Lemma of Borel-Cantelli, gives the inner part of Theorem 7.2.1: For all constants $\varepsilon > 0$ and $\mathcal{C}_I \geq \frac{1+\varepsilon}{K}$

$$\mathbb{P}\left[\exists n_\varepsilon : \forall n \geq n_\varepsilon, \mathcal{B}_{n-\mathcal{C}_I \ln n} \subset A_{V(n)}\right] = 1.$$

### 7.2.2 The Outer Bound

For the estimate of the outer error of Theorem 7.2.1, the upper bound in Proposition 7.1.11 is tighter than the analogous statement in [BB07, Proposition 2.3] because of non-amenability (see also [BB07, Remark 2.4]), which in turn leads to a smaller constant $\mathcal{C}_O$.

In this part of the proof we rely very much on the “onion structure” of the balls that was established in (7.8). We define random variables that count how many points of the IDLA-cluster lie in each layer of this ”onion”. For a fixed $n$ let

$$Z_p(j) = |A_j \cap \partial\mathcal{B}_{n+p}| = \sum_{i=1}^{j} \mathbf{1}_{\{X^i_{\sigma^i} \in \partial\mathcal{B}_{n+p}\}},$$

for $p \geq 1$, and

$$\nu_p(j) = \mathbb{E}[Z_p(j)].$$

Now, if $X^i_t$ adds to the cluster in a point of $\partial\mathcal{B}_{n+p+1}$, we know that it has always stayed in the cluster before that. Additionally, to arrive at a point in $\partial\mathcal{B}_{n+p+1}$, the random walk has to visit a point of $\partial\mathcal{B}_{n+p}$ first.

This means

$$\left[X^i_{\sigma^i} \in \partial\mathcal{B}_{n+p+1}\right] \subset \left[\exists x \in \partial\mathcal{B}_{n+p} \cap A_{i-1} : \tau^i_x < \infty\right],$$

which leads to

$$\begin{aligned}
\nu_{p+1}(j) &= \sum_{i=1}^{j} \mathbb{P}\left[X^i_{\sigma^i} \in \partial\mathcal{B}_{n+p+1}\right] \leq \sum_{i=1}^{j} \mathbb{P}\left[\exists x \in \partial\mathcal{B}_{n+p} \cap A_{i-1} : \tau^i_x < \infty\right] \\
&\leq \sum_{i=1}^{j} \sum_{x \in \partial\mathcal{B}_{n+p} \cap A_{i-1}} \mathbb{P}\left[\tau^i_x < \infty\right] = \sum_{i=1}^{j} \mathbb{E}\left[\sum_{x \in \partial\mathcal{B}_{n+p} \cap A_{i-1}} \mathbf{1}_{\{\tau^i_x < \infty\}}\right] \\
&\leq \sum_{i=1}^{j} \mathbb{E}\left[Z_p(i-1) \cdot \max_{x \in \partial\mathcal{B}_{n+p}} \mathbb{P}\left[\tau^i_x < \infty\right]\right].
\end{aligned}$$

All random walks are independent, therefore $X^j_t$ is also independent of $A_{j-1}$, and so

$$\begin{aligned}\nu_{p+1}(j) &\le \sum_{i=1}^{j} \nu_p(i-1) \cdot \max_{x\in\partial\mathcal{B}_{n+p}} \mathbb{P}[\tau_x^i < \infty] \\ &= \sum_{i=1}^{j} \nu_p(i-1) \cdot \max_{x\in\partial\mathcal{B}_{n+p}} F(0,x).\end{aligned}$$

If $x \in \partial\mathcal{B}_{n+p}$ this implies that $x \notin \mathcal{B}_{n+p}$, hence, by the definition of the balls $\mathcal{B}_n$, we have $F(0,x) < e^{-K(n+p)}$. Therefore

$$\nu_{p+1}(j) \le e^{-K(n+p)} \sum_{i=1}^{j} \nu_p(i-1). \tag{7.20}$$

Applying the inequality

$$\sum_{i=1}^{j} (i-1)^p \le \frac{j^{p+1}}{p+1}$$

and the trivial bound $\nu_1(j) \le j$ recursively to (7.20) gives

$$\nu_p(j) \le \exp\left(-Kn(p-1)\right) \exp\left(-K\frac{p(p-1)}{2}\right) \frac{j^p}{p!}.$$

Now we apply $p! \ge p^p e^{-p}$ and arrive at the final estimate

$$\nu_p(j) \le \exp\left(-Kn(p-1)\right) \exp\left(-K\frac{p(p-1)}{2}\right) e^p p^{-p} j^p. \tag{7.21}$$

The upper estimate of Proposition 7.1.11 gives, for $j = V(n)$

$$\begin{aligned}\nu_p(V(n)) &\le \exp(Kn) \exp\left(-K\frac{p(p-1)}{2}\right) e^p p^{-p} \mathcal{C}_u^p n^p \\ &= \exp\left(-K\left(\frac{p^2}{2} - n\right) + K\frac{p}{2} + p\right) n^p \left(\frac{\mathcal{C}_u}{p}\right)^p.\end{aligned}$$

For $p \ge \max\left\{\mathcal{C}_u, 6K^{-1}+3\right\}$ this leads to

$$\nu_p(V(n)) \le n^p \exp\left(-K\left(\frac{p^2}{3} - n\right)\right).$$

For $\mathcal{C}_O > \sqrt{3}$, $p \ge \mathcal{C}_O\sqrt{n}$ and some positive constant $c$,

$$\begin{aligned}\mathbb{P}[A_{V(n)} \not\subseteq \mathcal{B}_{n+\mathcal{C}_O\sqrt{n}}] &\le \mathbb{P}[Z_{\mathcal{C}_O\sqrt{n}}(V(n)) \ge 1] \\ &\le \nu_{\mathcal{C}_O\sqrt{n}}(V(n)) \\ &\le c \exp\left(-K\left(\frac{\mathcal{C}_O^2}{3} - 1\right) n\right),\end{aligned}$$

for $n \geq n_0$. Therefore

$$\sum_{n \geq n_0}^{\infty} \mathbb{P}[A_{V(n)} \nsubseteq \mathcal{B}_{n+\mathcal{C}_O\sqrt{n}}] < \infty,$$

and by the Lemma of Borel-Cantelli

$$\mathbb{P}\left[\exists n_0 : \, \forall n \geq n_0 : \, A_{V(n)} \subseteq \mathcal{B}_{n+\mathcal{C}_O\sqrt{n}}\right] = 1.$$

## Chapter 8

# Outlook

There are still several unsolved problems regarding internal growth models on the comb. The standard techniques to prove an outer bound, which where introduced by Lawler, Bramson and Griffeath [LBG92], and are also the basis for the proof in Section 7.2.2, do not work particularly well in the case of the comb. The method makes the assumption of uniform growth of the clusters, which on the comb is of course violated, since the cluster grows with rate $n^{1/3}$ in the direction of the $x$-axis, and with rate $n^{2/3}$ in the direction of the $y$-axis. Because of these problems, when we apply this method to the comb, using the harmonic measure of the sets $\mathcal{B}_n$ we obtained in Lemma 5.2.1, only a relatively weak upper bound of the form $A_n \subset \mathcal{B}_{n(1+\varepsilon n^{2/3})}$ can be obtained.

More promising in this context is probably the method Lawler used in his estimate of the fluctuations of the outer bound of IDLA on $\mathbb{Z}^d$ in [Law95]. There, in addition to some information on the harmonic measure of the sets $\mathcal{B}_n$, the quantities

$$nG_n(0,z) \quad \text{and} \quad \sum_{y\in\mathcal{B}_n} G_n(y,z)$$

are mainly used, where $G_n$ is again the Green function stopped at the first exit of $\mathcal{B}_n$. While the stopped Green function is not directly available on $\mathcal{C}_2$, in these special cases one can use the functions $g_n$ and $f_n$, as defined in (6.9) and (6.6), and the identities

$$nG_n(0,z) = \frac{d(z)}{2}\big(g_n(z) + f_n(z)\big),$$

and

$$\sum_{y\in\mathcal{B}_n} G_n(y,z) = \frac{d(z)}{2}\big(g_n(z) - f_n(z)\big).$$

Even though we obtained precise and explicit expressions for both $g_n$ and $f_n$ in Chapters 3 and 6, there remain several technical problems.

## 8.1 Internal growth models on the Sierpinski Carpet

In this section we describe a few simulation results for internal growth models on the graphical Sierpinski carpet in dimension 2. The graphical Sierpinski carpet is an infinite graph which is derived from the well know Sierpinski carpet, that is, the fractal which is created from the unit square in $\mathbb{R}^2$ by dividing it into 9 equal squares of which the one in the center is deleted. The same procedure is then repeated recursively to the remaining 8 squares.

The graphical Sierpinski carpet $S$ of dimension $d$ is defined as follows – see also Barlow and Bass [BB99].

**Definition 8.1.1.** *Let $S_0 = \mathbb{Z}_+^d$. Every vertex $v$ of $S_0$ can be written as $v = \big(v_1, v_2, \ldots, v_d\big)$ with all coordinates $v_i \geq 0$. We assume that every coordinate $v_i$ is written as an infinite ternary expansion, that is,*

$$v_i = \sum_{j=0}^{\infty} v_{ij} 3^j,$$

*where $v_{ij} \in \{0, 1, 2\}$ and $v_{i,j} \neq 0$ only for finitely many indices $j$. For each $k \geq 1$ set*

$$J_k = \Big\{\big(v_1, v_2, \ldots, v_d\big) \in S_0 : \ v_{ik} = 1 \text{ for all } 1 \leq i \leq d\Big\},$$

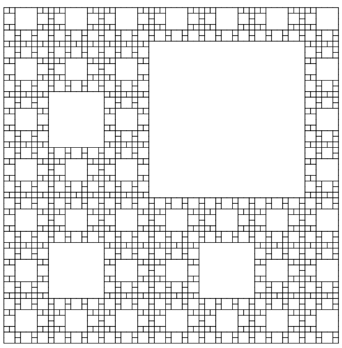


Figure 8.1: A finite piece of the graphical Sierpinski carpet in dimension 2.

*so that each $J_k$ is the union of disjoint cubes of side length $3^k$. Now define*

$$S_n = H_0 - \bigcup_{k=1}^{n} J_k \quad \textit{and} \quad S = \bigcap_{n=0}^{\infty} S_n.$$

*Then $S$ is the vertex set of the* graphical Sierpinsky carpet*, and the neighbourhood relation is inherited from $\mathbb{Z}_+^d$, that is, for $u, v \in S$ we have $u \sim v$ if and only if $|u - v| = 1$.*

Figure 8.1 shows a finite piece of the graphical Sierpinski carpet for $d = 2$.

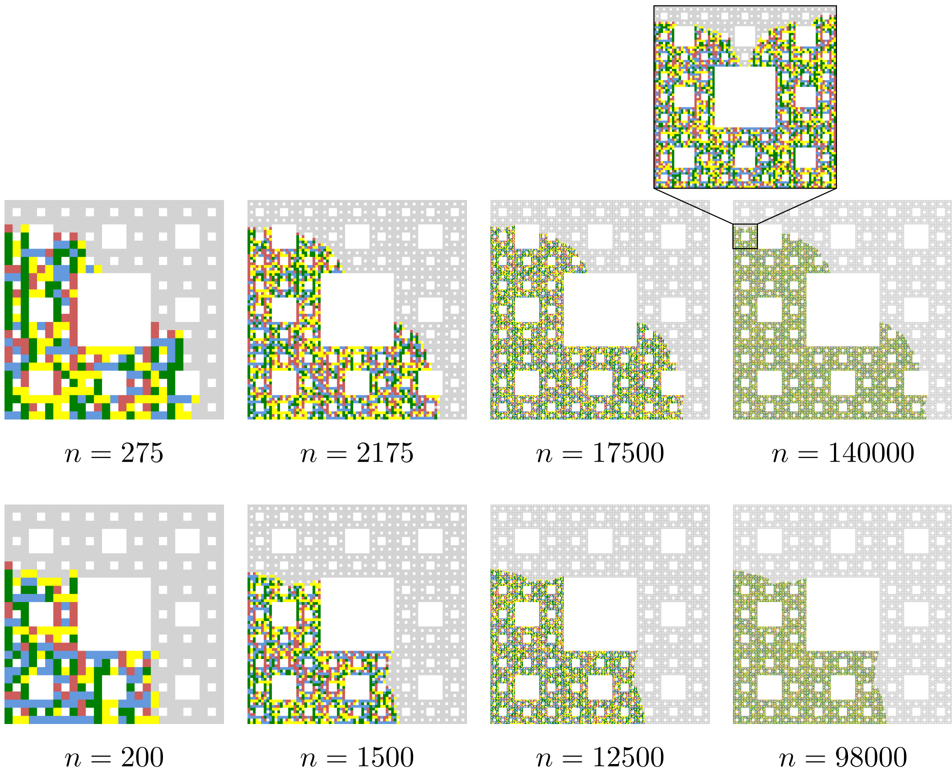


Figure 8.2: Evolution of the scaling limit of rotor-router clusters $R_n$ over 4 iterations of Sierpinski carpet. The scaling of each picture is $\frac{1}{3}$ of the picture to its left. The clusters in each row converge to two different scaling limits. The colouring represents the final states of the rotors.

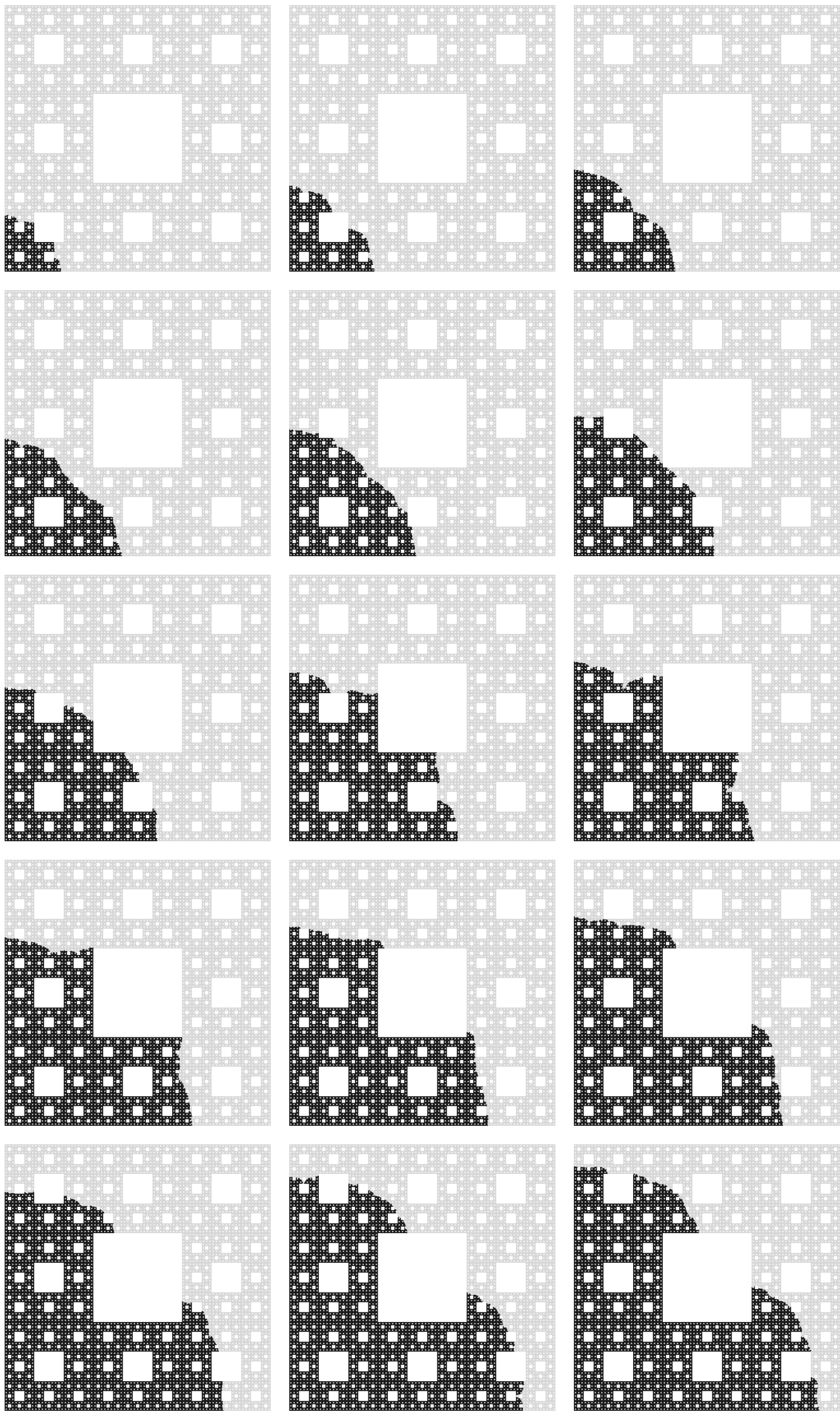


Figure 8.3: IDLA clusters on the Sierpinski carpet for 10000 up to 150000 particles.

Computer simulations of rotor-router aggregation and IDLA, with particles starting at the origin, exhibit an interesting behaviour. While, as expected, both growth models show similar behaviour, with IDLA having slightly bigger fluctuations, there does not seem to exist an unique scaling limit for the clusters. The simulations suggest that there is a whole family of scaling limits, dependending on how far the cluster has flown around the biggest hole in graph that is touched by the cluster. These scaling limits also seem to have a fractal boundary. Figure 8.2 shows two sequences of rotor-router clusters, which seem to converge to two different limiting sets. Figure 8.3 shows IDLA clusters on the Sierpinski carpet.

# Acknowledgements

I would like to thank my advisor Wolfgang Woess for his support without this work would not have been possible. Thanks also to Sébastian Blachère, Sara Brofferio, Steven Lalley, Lionel Levine, Yuval Peres, Ecaterina Sava, Florian Sobieczky and Christoph Temmel for many helpful discussions. I would like to thank Franz Lehner especially for helping me with the `FriCAS` computer system. Finally,t I have to thank my family and friends, without their support this work would not have been possible.

This work was partially supported by the Austrian Science Fund (FWF PROGRAM FWF-P19115-N18).

# Bibliography


[AG10a] A. Asselah and A. Gaudillière, *A note on fluctuations for internal diffusion limited aggregation*, (April 2010), `arXiv:1004.4665`.

[AG10b] A. Asselah and A. Gaudilliere, *From logarithmic to subdiffusive polynomial fluctuations for internal DLA and related growth models*, (September 2010), `arXiv:1009.2838`.

[AS92] N. Alon and J. H. Spencer, *The Probabilistic Method*, Wiley, New York, 1992.

[BB99] M. T. Barlow and R. F. Bass, Random walks on graphical Sierpinski carpets, in *Random walks and discrete potential theory*, edited by M. Picardello and W. Woess, CUP, 1999.

[BB07] S. Blachère and S. Brofferio, *Internal Diffusion Limited Aggregation on discrete groups having exponential growth*, Probability Theory and Related Fields **137 (3-4)** (2007), `arXiv:0507582`.

[Ber06] D. Bertacchi, *Asymptotic behaviour of the simple random walk on the 2-dimensional comb*, Electron. J. Probab. **11**, no. 45, 1184–1203 (electronic) (2006), `arXiv:0605718`.

[BHM08] S. Blachère, P. Haïssinsky and P. Mathieu, *Asymptotic entropy and Green speed for random walks on countable groups*, Ann. Probab. **36**(3), 1134–1152 (2008), `arXiv:0607467v2`.

[Bla00] S. Blachère, *Agrégation limitée par diffusion interne et temps de coupure sur les groupes discrets à croissance polynomiale*, PhD thesis, L'Université Paul Sabatier, 2000.

[Bla04] S. Blachère, Internal diffusion limited aggregation on discrete groups having polynomial growth, in *Random Walks and Geometry*, Proceedings (Erwin Schroedinger Insitute, Vienna 2001), de Gruyter, Berlin, 2004.

[BTW88] P. Bak, C. Tang and K. Wiesenfeld, *Self-organized criticality*, Phys. Rev. A **38**(1), 364–374 (Jul 1988).

[BZ03] D. Bertacchi and F. Zucca, *Uniform asymptotic estimates of transition probabilities on combs*, J. Aust. Math. Soc. **75**, 325–353 (2003), `arXiv:0008042v1`.

[CCFR09] E. Csáki, M. Csörgő, A. Földes and P. Révész, *Strong limit theorems for a simple random walk on the 2-dimensional comb*, Electron. J. Probab. **14**, no. 82, 2371–2390 (2009), `arXiv:0902.4369`.

[CCS09] M. Coornaert and T. Ceccherini Silberstein, A note on Laplace Operators on Groups, in *Limits of graphs in group theory and computer science*, edited by E. Press, Fundamental Sciences, pages 37–40, Lausanne Switzerland, 2009, CRC Press, 68Q80, 37B15, 20F65, 43A07.

[CS06] J. N. Cooper and J. Spencer, *Simulating a Random Walk with Constant Error*, Combinatorics, Probability and Computing **15**, 815–822 (2006), `arXiv:0402323v2`.

[DF91] P. Diaconis and W. Fulton, *A growth model, a game, an algebra, Lagrange inversion, and characteristic classes*, Rend. Sem. Mat. Univ. Pol. Torino **49**, 95–119 (1991).

[DF06] B. Doerr and T. Friedrich, Deterministic Random Walks on the Two-Dimensional Grid, in *ISAAC*, edited by T. Asano, volume 4288 of *Lecture Notes in Computer Science*, pages 474–483, Springer, 2006.

[DK86] J. Dodziuk and W. S. Kendall, *Combinatorial Laplacians and isoperimetric inequality*, From Local Times to Global Geometry, Control and Physics (ed. K. D. Elworthy), Pitman Res. Notes Math. Ser. **150**, 68–74 (1986).

[Dod84] J. Dodziuk, *Difference equations, isoperimetric inequality, and transience of certain random walks*, Trans. Amer. Math. Soc. **284**, 787–794 (1984).

[DTW03] I. Dumitriu, P. Tetali and P. Winkler, *On playing golf with two balls*, SIAM J. Discrete Math. **16**(4), 604–615 (electronic) (2003).

[FL] T. Friedrich and L. Levine, *Fast simulation of large-scale growth models*, `arXiv:1006.1003`.

[FS09] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge Univ. Press, Cambridge, 2009.

[Ger86] P. Gerl, *Natural spanning trees of $\mathbb{Z}^d$ are recurrent*, Discrete Mathematics **61**(2-3), 333 – 336 (1986).

[HLM+08] A. E. Holroyd, L. Levine, K. Mészáros, Y. Peres, J. Propp and D. B. Wilson, Chip-firing and rotor-routing on directed graphs, in *In and out of equilibrium. 2*, volume 60 of *Progr. Probab.*, pages 331–364, Birkhäuser, Basel, 2008.

[HP10] A. E. Holroyd and J. Propp, Rotor Walks and Markov Chains, in *Algorithmic Probability and Combinatorics*, edited by M. M. M. E. Lladser, Robert S. Maier and A. Rechnitzer, volume 520 of *Contemporary Mathematics*, pages 105–126, 2010.

[Hus08] W. Huss, *Internal diffusion-limited aggregation on non-amenable graphs*, Electron. Commun. Probab. **13**, 272–279 (2008).

[JLS10] D. Jerison, L. Levine and S. Sheffield, *Logarithmic fluctuations for internal DLA*, (October 2010), `arXiv:1010.2483`.

[KL] W. Kager and L. Levine, *Rotor-Router Aggregation on the layered square lattice*, `arXiv:1003.4017`.

[KL10] W. Kager and L. Levine, Diamond Aggregation, in *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 149, pages 351–372, 2010.

[Kle05] M. Kleber, *Goldbug variations*, Math. Intelligencer **27**(1), 55–63 (2005), `arXiv:0501497`.

[KP04] M. Krishnapur and Y. Peres, *Recurrent Graphs where Two Independent Random Walks Collide Finitely Often*, Electronic Communications in Probability **9**, 72–81 (2004), `arXiv:0406487`.

[Lam91] T. Y. Lam, *A first course in noncommutative rings*, volume 131 of *Graduate texts in mathematics*, Springer, New York, 1991.

[Law91] G. F. Lawler, *Intersections of Random Walks*, Birkäuser, Boston, 1991.

[Law95] G. F. Lawler, *Subdiffusive fluctuations for internal diffusion limited aggregation*, Ann. Probab. **23**, 71–86 (1995).

[LBG92] G. F. Lawler, M. Bramson and D. Griffeath, *Internal diffusion limited aggregation*, Ann. Probab. **20**, 2117–2140 (1992).

[Lev07] L. Levine, *Limit Theorems for Internal Aggregation Models*, PhD thesis, University of California, Berkeley, 2007.

[LL09] I. Landau and L. Levine, *The rotor-router model on regular trees*, J. Combin. Theory Ser. A **116**(2), 421–433 (2009), `arXiv:0712.3378`.

[LP] R. Lyons and Y. Peres, *Probabilty on trees and networks*, preprint.

[LP08] L. Levine and Y. Peres, *Spherical Asymptotics for the Rotor-Router Model in $\mathbb{Z}^d$*, Indiana Univ. Math. J **57**(1), 431–449 (2008), `arXiv:math/0503251`.

[LP09] L. Levine and Y. Peres, *Strong Spherical Asymptotics for Rotor-Router Aggregation and the Divisible Sandpile*, Potential Analysis **30**(1), 1–27 (2009), `arXiv:0704.0688v4`.

[LP10] L. Levine and Y. Peres, *Scaling Limits for Internal Aggregation Models with Multiple Sources*, Journal d'Analyse Mathématique (2010), `arXiv:0712.3378`, to appear.

[MM00] C. Moore and J. Machta, *Internal Diffusion-Limited Aggregation: Parallel Algorithms and Complexity*, J. Stat. Phys. **99**, 661–690 (2000), `arXiv:9909233`.

[MRZ01] R. Meester, F. Redig and D. Znamenski, *The Abelian Sandpile: a Mathematical Introduction*, Markov Processes and Related Fields **7**, 509–523 (2001), `arXiv:0301481`.

[NW02] T. Nagnibeda and W. Woess, *Random walks on trees with finitely many cone types*, J. Theoret. Probab. **15**(2), 383–422 (2002).

[PDDK96] V. B. Priezzhev, D. Dhar, A. Dhar and S. Krishnamurthy, *Eulerian Walkers as a Model of Self-Organized Criticality*, Phys. Rev. Lett. **77**(25), 5079–5082 (Dec 1996), `arXiv:9611019`.

[PPS98] A. M. Povolotsky, V. B. Priezzhev and R. R. Shcherbakov, *Dynamics of Eulerian walkers*, Phys. Rev. E **58**(5), 5449–5454 (Nov 1998), `arXiv:9802070`.

[She10] E. Shellef, *IDLA on the Supercritical Percolation Cluster*, Electronic Journal of Probability **15**, 723–740 (2010), `arXiv:0806.4771`.

[Spi76] F. Spitzer, *Principles of Random Walk*, Springer, 1976.

[WH86] G. H. Weiss and S. Havlin, *Some properties of a random walk on a comb structure*, Physica A: Statistical and Theoretical Physics **134**(2), 474–482 (1986).

[Woe00] W. Woess, *Random walks on infinite graphs and groups*, Cambridge Univ. Press, Cambridge, 2000.