SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

Internal Diffusion Limited Aggregation and Obstacle Problems

by

Kaj Börjeson

2011 - No 12

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET, 106 91 STOCKHOLM

Internal Diffusion Limited Aggregation and Obstacle Problems

Kaj Börjeson

Självständigt arbete i matematik 15 högskolepoäng, GN

Handledare: Henrik Shahgholian

2011

Internal Diffusion Limited Aggregation and obstacle problems

Kaj Börjeson

Abstract

We study the stochastic growth model known as Internal DLA which is a growth model in $\mathbb{Z}^d$. It features a cluster determined by adding new particles according to random walks. A particle travels from inside the cluster and stops when reaching the first point outside the region, adding that point to the region. We are mainly interested in the asymptotics of the occupied sets and present how this relates to PDE theory, in particular the obstacle problem and the Stefan Problem.

We also discuss some related models, the Divisible Sandpile, Rotor-Router Aggregation and the Classical Sandpile.

Contents

1 Introduction **3**
1.1 Internal DLA . . . 3
1.2 Divisible Sandpile . . . 4
1.3 Connection with obstacle problems . . . 5

2 Scaling limit shapes **8**
2.1 Shape theorem . . . 8
2.2 Stefan Problem . . . 10
2.2.1 Classical formulation . . . 10
2.2.2 Adapted formulation . . . 11
2.2.3 Continuous time IDLA . . . 11

3 IDLA with a single source **12**
3.1 Proof of the asymptotic shape . . . 12
3.1.1 Inner bound . . . 13
3.1.2 Outer bound . . . 17

4 Related Models **18**
4.1 Rotor-Router Aggregation . . . 18
4.2 Classical Sandpile . . . 20

5 Appendix **21**
5.1 The discrete and the continuous Laplace operator . . . 21
5.2 Discrete maximum-minimum principle . . . 23
5.3 Probabilistic concepts . . . 23
5.3.1 Expected value . . . 23
5.4 Stochastic processes . . . 23
5.4.1 Simple random walk . . . 24
5.4.2 Poisson process . . . 24
5.5 Green's function . . . 24
5.5.1 Continuous Green's function . . . 24
5.5.2 Discrete Green's function . . . 25
5.5.3 Green's function for a random walk in a ball . . . 27
5.6 Borel-Cantelli . . . 28
5.7 Sums of independent indicator variables . . . 28
5.8 Notations . . . 29

6 Acknowledgements **29**

1 Introduction

1.1 Internal DLA

Internal diffusion limited aggregation (IDLA) is a stochastic growth model in $\mathbb{Z}^d$. It was first introduced in [LBG] and later generalized in several directions, e.g. in [DF] with a more algebraic approach and in [GQ] with connections to partial differential equations. IDLA uses random walks to determine the growth of a cluster.

We consider systems of discrete particles in $\mathbb{Z}^d$ where $d \geq 2$ is the dimension.

Definition 1.1.1. *Let $f : \mathbb{Z}^d \to \mathbb{Z}_+$ have finite mass*

$$\sum_{x \in \mathbb{Z}^d} f(x) < \infty,$$

and call f a discrete mass configuration in $\mathbb{Z}^d$. We also say that there are $f(x)$ particles at the point x.

Definition 1.1.2. *Let $i \in \mathbb{Z}^d$. Then let T_i be a function from a discrete mass configuration m into another discrete mass configuration $T_i(m)$ defined as follows. Let $T_i(m) = m$ if $m(i) \leq 1$. For $m(i) > 1$, let X_{T_i} be a simple random walk (defined in Section 5.4.1) started at i and let $j \in \mathbb{Z}^d$ be the point where the random walk first hits the set $\{x | f(x) = 0\}$. Then*

$$T_i(m) = m - \delta_{ix} + \delta_{jx}$$

where δ is Kronecker's delta and $x \in \mathbb{Z}^d$. The function T_i is called a toppling at i. From now on, if we suppress the index or if we say that a function is a toppling function we mean that it is a toppling function at some (unspecified) i.

Intuitively this function takes a particle from the site i if there are more than one particle there and moves it, by a random walk, to the first empty site it hits.

Definition 1.1.3. *Let σ be a discrete mass configuration and let $\{T_{x_1}T_{x_2}, ..., T_{x_\alpha}\}$ be a set of toppling functions. If*

$$\sup_{x \in \mathbb{Z}^d} T_{x_\alpha}...T_{x_2}T_{x_1}(\sigma) \leq 1,$$

then

$$T_{x_\alpha}...T_{x_2}T_{x_1}(\sigma)$$

is called the final mass configuration obtained from the initial configuration σ (with respect to $\{T_{x_1}T_{x_2}, ..., T_{x_\alpha}\}$).

In [DF] it is shown that for any collection of toppling functions with the property that $\sup_{x\in\mathbb{Z}^d} T_{x_\alpha}...T_{x_2}T_{x_1}(\sigma) \leq 1$, the probability distribution of occupied domains is the same. That is, the order of the topplings does not matter. This is not hard to see by first considering the case of two particles. Suppose that the first particle hits an empty site and therefore the second particle does not stop at that site when it hits that site. But the particles are identical and from this follows that we cannot distinguish which of the particles continues to move. Therefore the result is identical. This enables us to formulate the definition of the final occupied domain and be sure that it is independent of the choice of toppling functions.

Definition 1.1.4. *Given a initial discrete mass configuration σ and any set of toppling functions $\{T_{x_1}T_{x_2}, ..., T_{x_\alpha}\}$ so that $\sup_{x\in\mathbb{Z}^d} T_{x_\alpha}...T_{x_2}T_{x_1}(\sigma) \leq 1$ we will call the discrete mass configuration*

$$T_{x_\alpha}...T_{x_2}T_{x_1}(\sigma)$$

a final mass configuration and denote it by ν. The set

$$\{x|\nu(x) = 1\}$$

is called the final domain of occupied sites.

The main problem is to describe how the final domain of occupied sites looks for a given initial mass configuration σ. For n particles starting at the origin, the final occupied domain will approach a Euclidean ball when n tends to infinity. This is Theorem 3.1.1. It is noteworthy since it is not a priori clear because of the structure of the lattice $\mathbb{Z}^d$.

1.2 Divisible Sandpile

In order to analyse IDLA with methods from potential theory and in particular the obstacle problem we will define a deterministic analogue with continuous mass. This is called the Divisible Sandpile and was introduced in [L]. The heuristic motivation for this model is that if there is equal chance for a particle to move to any neighbouring vertices (since we are considering simple random walks) then we instead consider dividing the mass equally among the neighbours. The definition for this model will be in the spirit of the discrete model.

Definition 1.2.1. *Let $f : \mathbb{Z}^d \to \mathbb{R}_+$ with*

$$\sum_{x\in\mathbb{Z}^d} f(x) < \infty,$$

and call f a continuous mass configuration in $\mathbb{Z}^d$. We also say that at the point x the amount of mass is $f(x)$.

Definition 1.2.2. *Let T_i be a function from a continuous mass configuration m into another continuous mass configuration $T_i(m)$. Let $T_i(m) = m$ if $m(i) \leq 1$. Otherwise let*

$$T_i(m) = m - (m(i) - 1)\delta_{ix} + \frac{1}{2d}\sum_{j \sim i}(m(i) - 1)\delta_{jx},$$

where δ is Kronecker's delta and $j \sim i$ means that j is adjacent to i (i.e $|i - j| = 1$). The function T_i is called a toppling at i.

The intuition behind this definition is that T_i takes all mass exceeding 1 at i and divides it equally between its $2d$ neighbours.

Definition 1.2.3. *Let σ be a continuous mass configuration and let $\{T_{x_1}T_{x_2}, ...\}$ be any sequence of toppling functions so that*

$$\lim_{\alpha \to \infty} \sup_{x \in \mathbb{Z}^d} T_{x_\alpha} ... T_{x_2} T_{x_1}(\sigma) \leq 1.$$

Then $\lim_{\alpha \to \infty} T_{x_\alpha} ... T_{x_2} T_{x_1}(\sigma)$ is called the final mass configuration obtained from the inital configuration σ (with respect to $\{T_{x_1}T_{x_2}, ...\}$).

As in the discrete case it can be shown that this is independent of the order of topplings (see [L]). Thus it makes sense to make the following definition.

Definition 1.2.4. *Let σ be an initial mass configuration and ν a mass configuration*

$$\lim_{\alpha \to \infty} T_{x_\alpha} ... T_{x_2} T_{x_1}(\sigma)$$

with the property that $\lim_{\alpha \to \infty} \sup_{x \in \mathbb{Z}^d} T_{x_\alpha} ... T_{x_2} T_{x_1}(\sigma) \leq 1$. Call ν a final mass configuration. The set

$$\{x | \nu(x) = 1\}$$

is called the final domain of fully occupied sites.

1.3 Connection with obstacle problems

We will reformulate the problem of finding the final occupied domain in terms that can be analysed by tools from potential analysis. The key step is the definition of a function called the odometer function. This is formulated for the Divisible Sandpile. We can then analyse the asymptotic behaviour of IDLA by the looking at divisible sandpile. A function similar to this appears naturally in the theory of obstacle problems. We follow the approach of [L].

Definition 1.3.1. *Let σ be an initial configuration of continuous mass and let ν be its final mass configuration after a sequence of topplings. Then for a specific site i there is an amount of mass emitted from i. We will define the odometer function $u(i)$ as the total amount emitted from i. More precise*

$$u(i) = \sum_{k|x_k=i} T_{x_{k-1}}...T_{x_2}T_{x_1}(\sigma(i)) - T_{x_k}...T_{x_2}T_{x_1}(\sigma(i)),$$

that is the sum of the differences in mass after every toppling at i.

The important feature of this function is the following lemma.

Lemma 1.3.1. *For the odometer function u we have the identity*

$$\Delta u = \nu - \sigma,$$

where σ is the initial mass distribution, ν is the final mass distribution and Δ is the discrete Laplace operator defined in section 5.1.

Proof. Since every site emits the same amount of mass to all of its neighbours, the mass recieved by the site x is $\frac{1}{2d}\sum_{y\sim x} u(y)$ where $y \sim x$ means that $|x-y| = 1$. We then have

$$\Delta u(i) = \frac{1}{2d}\sum_{i\sim j}(u(j) - u(i)) = -u(i) + \frac{1}{2d}\sum_{i\sim j} u(j) = \nu(i) - \sigma(i).$$

Note that our definition of the Laplace operator is scaled by dividing by a factor of $2d$ (see Section 5.1). □

Note that the odometer function only detects movement of mass. To obtain information about the set of nonempty sites we need some more information about the initial mass configuration. Consider for example an initial configuration where every site has either zero mass or mass greater than 1. Then every site in the final domain has either mass 1 or is a neighbour of a site with mass 1. In this case the odometer function is of great help when finding all occupied sites.

We have $\Delta u = 1 - \sigma$ in the final domain of occupied sites and $\Delta u \leq 1 - \sigma$ outside of the final domain. We will construct a function γ with $\Delta\gamma = \sigma - 1$. An example of such a function is

$$\gamma(x) = -|x|^2 - \sum_{y\in\mathbb{Z}^d} g_1(x,y)\sigma(y)$$

where $|x|$ is the usual Euclidean norm defined by $|x| = \sqrt{\sum_{i=1}^{d} |x_i|^2}$ and $g_1(x,y)$ is the discrete Green's function defined in section 5.5.2.

The next Lemma formulates the problem of finding the odometer function for the divisible sandpile as a discrete obstacle problem.

Lemma 1.3.2. *Let $\sigma : \mathbb{Z}^d \to \mathbb{R}_+$ be an initial continuous mass configuration for the divisible sandpile. Let*

$$S = \{f|\Delta f \leq 0, f \geq \gamma\},$$

with $\gamma(x) = -|x|^2 - \sum_{y\in\mathbb{Z}^d} g_1(x,y)\sigma(y)$. Then we have

$$u = s - \gamma$$

with u as the odometer function and

$$s = \inf_{f\in S} f.$$

Proof. First observe that by construction of γ, we have $\Delta(u+\gamma) \leq 0$. Note that since u is nonnegative it follows that $u+\gamma \geq s$. We see that $\Delta(u+\gamma) \geq 0$ and that $u+\gamma \geq \gamma$ so $u+\gamma \in S$. Hence $u+\gamma \geq s$. Now we prove the reverse inequality. For $f \in S$

$$\Delta(f - \gamma - u) \leq 0$$

on the final domain of occupied sites since $\Delta u = 1 - \sigma$ on that domain. We also see that it is nonnegative outside the final domain of occupied sites. By Theorem 5.2.1 we conclude that $f - \gamma - u$ is nonnegative everywhere. Thus $u + \gamma \leq s$ and the result follows. □

This Lemma allows us to define a continuous analogue of the divisible sandpile. It is defined by taking the dual view by constructing a continuous odometer function which then determines a set where we say that there is mass 1. One has to be careful since the odometer function does not determine the final occupied cluster uniquely. However if the initial configuration is sufficiently nice (for example demanding that every x satisfy $\sigma(x) = 0$ or $\sigma(x) \geq 1 + \epsilon$ for some fixed uniform ϵ) the final domain is determined uniquely except at the boundary. But the situation at the boundary will in one sense be easier since in the continuous case the measure of the boundary will be small and therefore the mass located at the boundary is negligible.

Definition 1.3.2. *Let $\sigma : \mathbb{R}^d \to \mathbb{R}_+$ be a function with bounded support. Then define the continuous odometer function as*

$$u = s - \gamma,$$

where

$$\gamma(x) = -|x|^2 - \int_{\mathbb{R}^d} g(x,y)\sigma(y)dy,$$

with $g(x,y)$ as defined in section 5.5.1 and

$$s(x) = \inf\{f(x)|\Delta f \leq 0 \text{ and } f \geq \gamma\}.$$

This is called the obstacle problem for the obstacle γ. This is a well-studied problem; for a reference see [F]. This interpretation will allow us to analyse the discrete models by using the continuous model which is easier to control. The proof of the Theorem 2.1.1 uses the basic structure of the proof of Theorem 3.1.1 but with one crucial difference, the domain is not the circle but is decided by the set of occupied sites of the Divisible Sandpile. We will point out in the proof where this happens. There is also a lot of regularity arguments which is why we don't present the full proof of Theorem 2.1.1 here and only the conceptually easier Theorem 3.1.1.

2 Scaling limit shapes

2.1 Shape theorem

For IDLA with a single source the asymptotic shape of the domain of occupied sites is a Euclidean ball. This is Theorem 3.1.1 and was proved in [LBG] and we present a proof of it in section 3. A natural question to ask is whether there exists asymptotic domains for more general starting configurations. We will introduce a scaling that reduces the size of the mesh and compensates by adding more particles. In the limit where the mesh size approaches zero we have limiting shapes for many starting configurations. Theorem 2.1.1 describes the situation. We again follow the approach of [L].

Definition 2.1.1. *Let $B(x, \epsilon)$ denote the open ball of radius ϵ around the point x. Then let A_ϵ be defined as*

$$A_\epsilon = \{x \in A | B(x, \epsilon) \subset A\}$$

and A^ϵ as

$$A^\epsilon = \{x \in \mathbb{R}^d | B(x, \epsilon) \cap A \neq \emptyset\}.$$

A_ϵ is called the inner ϵ-neighbourhood and A^ϵ the outer ϵ-neighbourhood.

Definition 2.1.2. *Let $A(n) \subset \delta_n \mathbb{Z}^d$ and $A \subset \mathbb{R}^d$ where δ_n is a sequence of positive real numbers approaching 0. We will say that the sequence $A(n)$ converges to A if for every $\epsilon > 0$ we have*

$$A_\epsilon \cap \delta_n \mathbb{Z}^d \subset A(n) \subset A^\epsilon$$

for sufficiently large n. We also write $A(n) \to A$.

Our next definition defines our scaling. It is done by considering a mass configuration defined in $\mathbb{R}^d$ and for every δ_n define a discrete mass configuration in $\mathbb{Z}^d$ so that the mass at each point is approximately the mean mass around the point.

Definition 2.1.3. *Let $\sigma : \mathbb{R}^d \to \mathbb{Z}_+$ be a locally integrable function. We define the discretization of σ at x with respect to the grid $\delta_n\mathbb{Z}^d$ as*

$$\sigma_n(x) = \left\lfloor \delta_n^{-d} \int_{x^\square} \sigma(y)dy \right\rfloor$$

where $x^\square$ denotes the set $x + [-\delta_n/2, \delta_n/2]^d$ and $\lfloor a \rfloor$ denotes the closest integer to a (rounding down).

Definition 2.1.4. *Let σ_n be an initial mass configuration in the grid $\delta_n\mathbb{Z}^d$. The final domain of occupied sites for the divisible sandpile with starting mass σ_n is denoted by D_n and the final domain of occupied sites for Internal DLA by I_n.*

Definition 2.1.5. *Let $\sigma : \mathbb{R}^d \to \mathbb{R}_+$ be a function with bounded support. Define*

$$D = \{x \in \mathbb{Z}^d | \gamma(x) < s(x)\},$$

where

$$\gamma(x) = -|x|^2 - \int_{\mathbb{R}^d} g(x,y)\sigma(y)dy,$$

and

$$s(x) = \inf\{f(x) | \Delta f \leq 0 \text{ and } f \geq \gamma\}.$$

Theorem 2.1.1. *Let $d \geq 2$ and $\sigma : \mathbb{R}^d \to \mathbb{Z}_+$ be a bounded function that is almost everywhere continuous. Suppose there exist an open set $\Omega \subset \mathbb{R}^d$ such that $\{\sigma \geq 1\} = \bar{\Omega}$. Then as $n \to \infty$,*

$$D_n \to D \cup \Omega.$$

Furthermore, if $\delta_n \leq 1/\log(n)$,

$$I_n \to D \cup \Omega$$

with probability one.

Proof. Due to the amount of technicalities we will not prove this here, however we prove a special case in Section 3. For the full proof, see [L]. □

This means that for sufficiently nice initial configurations the scaling limit of the divisible sandpile and IDLA is determined by the solution of a continuous obstacle problem. For example one can see deduce that the scaling limit of IDLA from a single site must be circular because the continuous obstacle problem is rotationally invariant. Furthermore, by invoking results of potential theory and complex analysis and our models relation to quadrature domains we can determine the scaling limit of IDLA with n sources when $d = 2$. The boundary is then an algebraic curve of degree $2n$ ([L]).

Our next step is to relate this to a free boundary problem which will also provide us with a physical interpretation of our model.

2.2 Stefan Problem

We will see that IDLA acts as a discretized version of a free boundary problem known as the Stefan Problem if we consider the particle system as depending on time. We follow the approach of [GQ]. By considering the special case $t = \infty$ with no heat source we have an intuition for the situation in Theorem 2.1.1.

2.2.1 Classical formulation

Definition 2.2.1. *Let $D \subset \mathbb{R}^d$ be a domain defined by $D_1 \setminus D_0$ for two solid domains D_0 and D_1 with $\bar{D}_0 \subseteq D_1^\circ$. Set $\Gamma_i = \partial D_i$ $(i = 0, 1)$ and assume that Γ_i is a two times differentiable closed hypersurface and assume that there is a function $g(x,t)$ defined on Γ_0, a function $h(x)$ on D and a constant $k > 0$. A pair of functions $\theta(x,t)$ and $s(x)$ is called a solution of the classical Stefan problem if they satisfy*

$$s(x) = 0 \text{ if and only if } x \in \Gamma_1,$$

$$\theta(x,0) = h(x) \text{ for } x \in D,$$

$$\theta(x,t) = g(x,t) \text{ for } x \in \Gamma_0,$$

$$\theta(x,t) = 0 \text{ for } x \text{ such that } s(x) = t,$$

$$\nabla_x \theta(x,t) \cdot \nabla_x s(x) = -k \text{ for } x \text{ such that } s(x) = t,$$

and

$$\partial_t \theta(x,t) - \Delta_x \theta(x,t) = 0 \text{ for } x \text{ such that } s(x) < t \text{ or } x \in D.$$

This problem enjoys existence and uniqueness of solutions for some natural regularity assumptions. However, we shall not consider this problem in detail and refer the reader to [F].

To provide a bit of intuition, we will describe a physical situation where these equations occur. Suppose we have a system of water, ice and a heat source. Let D be the set occupied by water at $t = 0$. Let D_0 be a heat source. We shall not consider the situation inside the heat source and model it only by the fixed temperature $g(x,t)$ on the surface Γ_0. The set D_1^c is considered to be occupied by ice at $t = 0$. The initial temperature of the water is given by the function $h(x)$ and the ice will be kept at 0. Now consider the melting of ice and let $\theta(x,t)$ be the temperature at point x at time t. The moving interface between water and ice is described by the function $s(x)$ such that the curve defined by the solutions to the equation $s(x) - t = 0$ is the boundary at time t. The region where $s(x) < t$ is occupied by water and the region where $s(x) > t$ is occupied by ice at time t. The constant k is a physical constant related to the specific melting heat of ice.

Now because we would like the free boundary at $t = 0$ to be equal to Γ_1 we get the restriction that $s(x) = 0$ if and only if $x \in \Gamma_1$. We have $\theta(x,0) =$

$h(x)$ for $x \in D$ and $\theta(x,t) = g(x,t)$ for $x \in \Gamma_0$ by the definition of h and g. We have $\theta(x,t) = 0$ for x such that $s(x) = t$ by continuity of temperature at the free boundary. The condition that $\nabla_x \theta(x,t) \cdot \nabla_x s(x) = -k$ for x such that $s(x) = t$ is due to the physics of melting, in short the speed of melting is proportional to the temperature of the surrounding water and the specific melting heat of ice. The last condition that $\partial_t \theta(x,t) - \Delta_x \theta(x,t) = 0$ for x such that $s(x) < t$ or $x \in D$ is the ordinary heat equation for all points in the interior of the water region.

2.2.2 Adapted formulation

Our adapted formulation is almost the same as the classical with the difference that we model the "heat source" with points. For that we will use the Dirac delta function and obtain the equation $\partial_t \theta(x,t) - \Delta_x \theta(x,t) - \sum_{i=1}^{n} c_i \delta_{x_i} = 0$. Specifically we have the following setup.

Definition 2.2.2. *Let $D \subset \mathbb{R}^n$ be a solid domain. Then define $\Gamma = \partial D$ and assume that Γ is a two times differentiable closed hypersurface. Furthermore assume that there is a finite collection of points $\{x_i\}_i$ inside D together with positive constants $\{c_i\}_i$, a function $h(x)$ on D and a constant $k > 0$. A pair of functions $\theta(x,t)$ and $s(x)$ is called a solution of the adapted Stefan problem if they satisfy*

$$s(x) = 0 \text{ if and only if } x \in \Gamma,$$

$$\theta(x,0) = h(x) \text{for } x \in D,$$
$$\theta(x,t) = 0 \text{ for } x \text{ such that } s(x) = t,$$
$$\nabla_x \theta(x,t) \cdot \nabla_x s(x) = -k \text{ for } x \text{ such that } s(x) = t,$$

and

$$\partial_t \theta(x,t) - \Delta_x \theta(x,t) - \sum_{i=1}^{n} c_i \delta_{x_i} = 0 \text{ for } x \text{ such that } s(x) < t \text{ or } x \in D,$$

where δ_{x_i} denotes the Dirac delta at x_i.

2.2.3 Continuous time IDLA

We will now consider a model of IDLA where the particles move simultaneously in continuous time. It features both an initial configuration of particles and point sources where new particles are created. Both the jumping times and the particle creation will be modeled using Poisson processes (defined in Section 5.4.2).

Definition 2.2.3. *Assume the initial setup of particles is the same as in Definition 2.1.3 with grid size δ_n. Furthermore suppose there is a finite set of point sources $\{x_i\}_i$ with corresponding intensities $\{c_i\}_i$. The generalized IDLA model is the process occuring if every particle has a probability to jump to a neighbouring vertex according to a Poisson process with intensity δ_n^{-2} and every point source x_k creates a new particle with a certain probability according to a Poisson process with intensity $c_k\delta_n^{-d}$.*

By considering the definition without heat sources and look at the configuration at $t = \infty$ this will correspond to the earlier definition of IDLA. This is because every particle will stop with probability 1 and the order of the particles does not matter. We will not state the exact conditions and type of convergence of this model, the details are found in [GQ].

Theorem 2.2.1. *As $\delta_n \to 0$ the density of particles in the generalized IDLA model converges weakly to the unique solution $\theta(x, t)$ of the adapted Stefan problem with corresponding initial conditions and point sources.*

Here $h(x)$ of the Stefan problem is set to the initial density of the particles in the IDLA model. The constant k in the Stefan problem is the amount of particles every site absorbs before becoming full so according to our definition of IDLA the constant will be $k = 1$.

Proof. See [GQ]. □

3 IDLA with a single source

3.1 Proof of the asymptotic shape

In this section we will present a proof of the circular asymptotics of IDLA with a single source. The purpose of this is mainly to showcase the techniques used when proving results about IDLA. The proof follows [LBG] which is one of the first articles on the subject. We will not prove all lemmas but focus on the heart of the matter. The lemmas used here can be found in the Appendix. In this section let all balls be lattice balls, i.e. the usual Euclidean ball intersected with $\mathbb{Z}^d$.

Theorem 3.1.1. *For any $\epsilon > 0$ we have*

$$B(0, n(1-\epsilon)) \subset D_{\lfloor \omega_d n^d \rfloor} \subset B(0, n(1+\epsilon)),$$

for large enough n with probability 1 where $D_{\lfloor \omega_d n^d \rfloor}$ is the IDLA cluster obtained by starting with $\lfloor \omega_d n^d \rfloor$ particles at the origin. Here ω_d is the volume of the d-dimensional unit ball and $\lfloor a \rfloor$ denotes the closest integer to a (rounding down).

Proof. We will prove the theorem by first showing that for any $\epsilon > 0$ and large enough n,

$$B(0, n(1-\epsilon)) \subset D_{\lfloor \omega_d n^d (1+\epsilon) \rfloor}$$

with probability 1 and then using this to prove that for any $\epsilon > 0$ and large enough n,

$$D_{\lfloor \omega_d n^d \rfloor} \subset B(0, n(1 + K_d \epsilon^{1/d}))$$

with probability 1 for some constant K_d and by this deduce the result. That the first inclusion is sufficient can be seen for example by the change of variables $n' = n(1+\epsilon)^{1/d}$ and observe that the quantity in front of n' in the left hand side also approaches 1 from below when $\epsilon \to 0^+$.

3.1.1 Inner bound

To prove that for any $\epsilon > 0$ and large enough n,

$$B(0, n(1-\epsilon)) \subset D_{\lfloor \omega_d n^d (1+\epsilon) \rfloor}$$

with probability 1, we will introduce some quantities that we will estimate. First, consider the random walks $X^i(t)$ ($X^i(t)$ denotes the i:th walk) determining the cluster but we let them evolve forever, even after they leave the cluster. Then we define the following random stopping times

$$\sigma_i = \min\{t | X^i(t) \notin D_{i-1}\},$$

$$\tau_z^i = \min\{t | X^i(t) = z\},$$

and

$$\tau_n^i = \min\{t | X^i(t) \notin B(0, n)\}.$$

These can be interpreted as follows, σ_i is the time of the first exit of the i:th particle from the cluster, τ_z^i is the first hitting time of z for the i:th particle and τ_n^i is the first exit time of $B(0, n)$ for the i:th particle.

To prove $B(0, n(1-\epsilon)) \subset D_{\lfloor \omega_d n^d (1+\epsilon) \rfloor}$ for large enough n we will use the following lemma.

Lemma 3.1.1. *If*

$$\sum_{n=1}^{\infty} \sum_{z \in B(0, n(1-\epsilon))} \mathbb{P}(\tau_z^i > \sigma_i \text{ for all } i \leq \lfloor \omega_d n^d (1+\epsilon) \rfloor) < \infty. \tag{1}$$

then $B(0, n(1-\epsilon)) \subset D_{\lfloor \omega_d n^d (1+\epsilon) \rfloor}$ *for large enough* n.

Proof. First note that

$$B(0, n(1-\epsilon)) \subset D_{\lfloor \omega_d n^d (1+\epsilon) \rfloor},$$

for large enough n is equivalent to that

$$\limsup_{n\to\infty} \mathbb{P}(B(0, n(1-\epsilon)) \subset D_{\lfloor \omega_d n^d (1+\epsilon) \rfloor}) = 1,$$

which in turn is equivalent to that

$$\limsup_{n\to\infty} \mathbb{P}(B(0, n(1-\epsilon)) \not\subseteq D_{\lfloor \omega_d n^d (1+\epsilon) \rfloor}) = 0.$$

Now

$$B(0, n(1-\epsilon)) \not\subseteq D_{\lfloor \omega_d n^d (1+\epsilon) \rfloor}$$

can be rephrased as

$$\tau_z^i > \sigma_i \text{ for all } i \leq \lfloor \omega_d n^d (1+\epsilon) \rfloor.$$

for some $z \in B(0, n(1-\epsilon))$.

Now applying the Borel-Cantelli (Theorem 5.6.1) yields the result.

□

To estimate the terms in this sum we introduce the following random variables

$$N = \sum_{i=1}^{\lfloor \omega_d n^d (1+\epsilon) \rfloor} 1_{\{\tau_z^i < \sigma^i\}},$$

$$M = \sum_{i=1}^{\lfloor \omega_d n^d (1+\epsilon) \rfloor} 1_{\{\tau_z^i < \tau_n^i\}},$$

and

$$L = \sum_{i=1}^{\lfloor \omega_d n^d (1+\epsilon) \rfloor} 1_{\{\sigma^i \leq \tau_z^i < \tau_n^i\}}.$$

These can be interpreted as follows, N is the number of particles hitting z before stopping, M is the number of walks hitting z before exiting the ball $B(0, n)$ and L is the number of walks visiting z before leaving the ball $B(0, n)$ but after the particle stops. Now we can see that

$$\mathbb{P}(\tau_z^i > \sigma_i \text{ for all } i \leq \lfloor \omega_d n^d (1+\epsilon) \rfloor) = \mathbb{P}(N = 0)$$

and that

$$N \geq M - L.$$

Lemma 3.1.2.
$$\mathbb{P}(N = 0) \leq \mathbb{P}(M \leq a) + \mathbb{P}(L \geq a)$$
for any given a.

Proof. First we show that

$$\mathbb{P}(N = 0) \leq \mathbb{P}(M \leq a \text{ or } L \geq a)$$

for any given a. This can be seen by considering cases. Suppose $M \leq a$, then we are done. Suppose instead that $M > a$, then $L > a$ since $L \geq M$. Now the Lemma follows from elementary probabilistic considerations. $\square$

Our strategy is now to estimate these terms and show that for some a they can be made small. To estimate these we are going to use the large deviations estimate for sums of independent variables (see Theorem 5.7.1). M is a sum of independent variables, however L is not a sum of independent variables and to estimate this we will enlarge the index set. Every summand corresponds to some stopping time σ^i and some point z. If we include all points of $B(0, n)$ and consider the post-t_z^i random walks we get the random variable

$$\bar{L} = \sum_{y \in B(0,n)} 1^y_{\{\tau_z^i < \tau_n^i\}},$$

which is a sum of independent variables. The point is that since $L \leq \bar{L}$ we can estimate $\mathbb{P}(\bar{L} \geq a)$ instead of $\mathbb{P}(L \geq a)$. So we have the following inequality

$$\mathbb{P}(\tau_z^i > \sigma_i \text{ for all } i \leq \lfloor \omega_d n^d (1 + \epsilon) \rfloor) \leq \mathbb{P}(M \leq a) + \mathbb{P}(\bar{L} \geq a),$$

for any given a.

Lemma 3.1.3. *For fixed $\epsilon > 0$ and for large enough n,*

$$\mathbb{P}(\bar{L} \geq (1 + \epsilon/4)\mathbb{E}[\bar{L}]) \leq \exp(-c_d n),$$

and
$$\mathbb{P}(M \leq (1 + \epsilon/4)\mathbb{E}[\bar{L}]) \leq \exp(-c_d n).$$

By putting $a = (1 + \epsilon/4)\mathbb{E}[\bar{L}])$ the sum (1) converges since every term is dominated by a decreasing exponential function.

Proof. First note that

$$\mathbb{E}[M] = \lfloor \omega_d n^d (1 + \epsilon) \rfloor \mathbb{P}(\tau_z < \tau_n),$$

and

$$\mathbb{E}[\bar{L}] = \sum_{y \in B(0,n)} \mathbb{P}_y(\tau_z < \tau_n),$$

where $\mathbb{P}_y(\tau_z < \tau_n)$ is the probability that a random walk starting at y hits z before leaving $B(0,n)$.

By standard results about Green's functions (for example [LL]) we can write

$$\mathbb{P}_y(\tau_z < \tau_n) = \frac{g_{B(0,n)}(y,z)}{g_{B(0,n)}(z,z)},$$

where $g_{B(0,n)}$ is the Green's function for a random walk stopped when exiting $B(0,n)$ see (5.5.3). By using these identities and Theorem 5.5.5 we get

$$\mathbb{E}[M] = \lfloor \omega_d n^d (1+\epsilon) \rfloor \frac{g_{B(0,n)}(0,z)}{g_{B(0,n)}(z,z)} \geq (1+\epsilon/2) \sum_{y \in B(0,n)} \frac{g_{B(0,n)}(y,z)}{g_{B(0,n)}(z,z)} = (1+\epsilon/2)\mathbb{E}[\bar{L}],$$

which relates M and $\bar{L}$. We can also rewrite

$$\mathbb{E}[\bar{L}] = \sum_{y \in B(0,n)} \frac{g_{B(0,n)}(y,z)}{g_{B(0,n)}(z,z)} = \frac{\mathbb{E}_z[\tau_n]}{g_{B(0,n)}(z,z)}.$$

Now by Theorem 5.5.5 we obtain, for $d = 2$

$$\mathbb{E}[\bar{L}] \geq \beta_2 n^2 / \log n,$$

and for $d \geq 3$

$$\mathbb{E}[\bar{L}] \geq \beta_d n^2,$$

for large enough n and for suitable positive constants β_d which depend on ϵ. By our estimate $\mathbb{E}[M] \geq (1+\epsilon/2)\mathbb{E}[\bar{L}]$ we have the same lower bound for $\mathbb{E}[M]$. Now since M and L are independent sums we can apply Theorem 5.7.1 with $\gamma = 1/3$ to obtain

$$\mathbb{P}(\bar{L} \geq \mathbb{E}[\bar{L}] + \mathbb{E}[\bar{L}]^{5/6}) \leq 2 \exp\Big(\frac{-\mathbb{E}[\bar{L}]^{2/3}}{4}\Big) \leq \exp(-c_d n),$$

and

$$\mathbb{P}(M \leq \mathbb{E}[M] - \mathbb{E}[M]^{5/6}) \leq 2 \exp\Big(\frac{-\mathbb{E}[M]^{2/3}}{4}\Big) \leq \exp(-c_d n).$$

Now by the relation $\mathbb{E}[M] \geq (1+\epsilon/2)\mathbb{E}[\bar{L}]$ the lemma follows.

□

Remark 3.1.1. *It can be interesting to note that almost the entire argument go through in the case of more general starting configurations. The one thing which fails is also the only argument where the domain as a ball appears explicitly, namely the lemma that for n sufficiently large and $z \in B(0, n(1-\epsilon))$ we have*

$$\sum_{y \in B(0,n)} g_{B(0,n)}(y, z) \leq \omega_d n^d g_{B(0,n)}(0, z).$$

This is saved (modulo some technicalities) by replacing the left hand side by

$$\sum_{y \in D} g_D(y, z),$$

where D is the final occupied domain of the Divisible Sandpile and replacing the right hand side by

$$\sum_{y \in D} \sigma(y) g_D(y, z),$$

where σ is the initial starting configuration. But the domain of the Divisible Sandpile is close to the domain obtained from the continuous obstacle problem (the support of the odometer function) in the limit for sufficiently nice starting configurations. This suggests that the limit shape is the support of the corresponding odometer function and is made precise in Theorem 2.1.1.

3.1.2 Outer bound

To obtain the outer bound we are going to use the inner estimate in an essential way. The point is that most of the particles stay to fill up $B(0, n(1-\epsilon))$ with only a small portion leaving $B(0, n)$. The next thing to check is that the particles exiting $B(0, n)$ are fairly spread out and do not create tentacles or other thin objects. To begin we define the event F as when the following estimate holds,

$$|D_{\lfloor \omega_d n^d \rfloor} \cap B(0, n)^c| < C_0 \epsilon n^d$$

with C_0 chosen so that

$$\mathbb{P}(F) \geq 1 - \epsilon,$$

by the inner bound. To control the particles leaving $B(0, n)$ we define $\bar{D}(i)$ as the shape after the ith particle leaving $B(0, n)$ has stopped. Define $Z_k(i)$ as

$$Z_k(i) = |\bar{D}(i) \cap S_{\lfloor 1 + K_d \epsilon^{1/d} \rfloor + 1 + k}|,$$

where S_j is the jth shell, namely $S_j = \{x \text{ so that } j \leq |x| < j + 1\}$. $Z_k(i)$ is interpreted as the amount of particles in the k:th outer shell after i particles has left the ball. We would like to prove

$$\mathbb{P}(\limsup_n \{D_{\lfloor \omega_d n^d \rfloor} \not\subseteq B(0, n(1 + K_d \epsilon^{1/d}))\} \cap F) = 0$$

as this together with the inner bound will imply our result. By Theorem 5.6.1 it is enough to prove that

$$\sum_n \mathbb{P}(D_{\lfloor \omega_d n^d \rfloor} \not\subseteq B(0, n(1 + K_d \epsilon^{1/d})), F) < \infty.$$

We will therefore estimate these terms as

$$\mathbb{P}(D_{\lfloor \omega_d n^d \rfloor} \not\subseteq B(0, n(1 + K_d \epsilon^{1/d})), F) \leq \mathbb{P}(Z_{n'}(\lfloor C_0 \epsilon n^d \rfloor \geq 1)) \leq \mathbb{E}[Z_{n'}(\lfloor C_0 \epsilon n^d \rfloor)]$$

for large enough n and $n' = \lfloor ((K-1)\epsilon^{1/d} n) \rfloor$ and use the following lemma:

Lemma 3.1.4. *For some constant C_1 we have*

$$\mathbb{E}[Z_k(j)] \leq n^{d-1} \big(C_1 \frac{j}{k} \epsilon^{(1-d/d)} n^{1-d}\big)^k.$$

Proof. We will not prove this lemma here. The technique used is to exploit that to reach an outer shell a particle must pass through all inner shells. This leads to a recursive estimate. The full proof is found in [LBG]. □

Using this lemma we see that

$$\mathbb{P}(D_{\lfloor \omega_d n^d \rfloor} \not\subseteq B(0, n(1 + K_d \epsilon^{1/d})), F) \leq n^{d-1} (\frac{C_1}{K})^{n'}$$

for large enough n and K. Now letting $K > C_0$ this will be bounded by $\exp(\alpha n)$ for some positive α. This exponential tail clearly implies the convergence of

$$\sum_n \mathbb{P}(D_{\lfloor \omega_d n^d \rfloor} \not\subseteq B(0, n(1 + K_d \epsilon^{1/d})), F),$$

which in turn implies the theorem since we can choose ϵ arbitrarily.

□

4 Related Models

4.1 Rotor-Router Aggregation

We will describe a model that is related to IDLA, but has the advantage that it is deterministic. It was introduced in [PDDK] under the name Eulerian Walkers. Suppose we have an initial discrete mass configuration σ as in Definition 1.1.1. Suppose moreover that we have a function $r(x) : \mathbb{Z}^d \to \{0, 1, ..., 2d-2, 2d-1\}$ assigning a "direction" to every point in $\mathbb{Z}^d$ where we have ordered the $2d$ neighbours of a point. In 2 dimensions we can imagine the directions as North(0), West(1), South(2) and East(3). Now the toppling of a particle at x is defined by moving a particle from x to the site where the function r points at and afterwards changing the value $r(x)$ to $r(x) + 1(\text{mod } 2d)$. This corresponds to changing the direction at x to the next neighbour. It can be shown that the final domain does not depend on the ordering of the topplings. This model enjoys results in the same sense as those for IDLA.

Definition 4.1.1. *With σ and σ_n as in Definition 2.1.3, define R_n to be the final domain of occupied sites for the Rotor-Router model with some fixed initial rotor configuration r.*

Theorem 4.1.1 ([L])**.** *Let $d \geq 2$. Let $\sigma : \mathbb{R}^d \to \mathbb{Z}_+$ be a bounded function that is continuous almost everywhere. Suppose there exist an open set $\Omega \subset \mathbb{R}^d$ such that $\{\sigma \geq 1\} = \bar{\Omega}$. Then as $n \to \infty$,*

$$R_n \to D \cup \Omega,$$

for any initial rotor configuration r and where D is as in Definition 2.1.5.

Numerical simulations suggest that the Rotor-Router model converge much faster and with very few anomalies to the shape of the divisible sandpile than IDLA does (see [C] for a collection of information obtained from simulations). There are several conjectures that seem reasonable, mostly concerning the Rotor-Router shape for n particles, starting at the origin with all rotors pointing in the same direction. For example, is the shape simply connected for all n? Is it convex for all n? Does the difference in radius of the smallest ball containing the shape and the largest ball contained in the shape stay bounded? Does the difference go to zero?

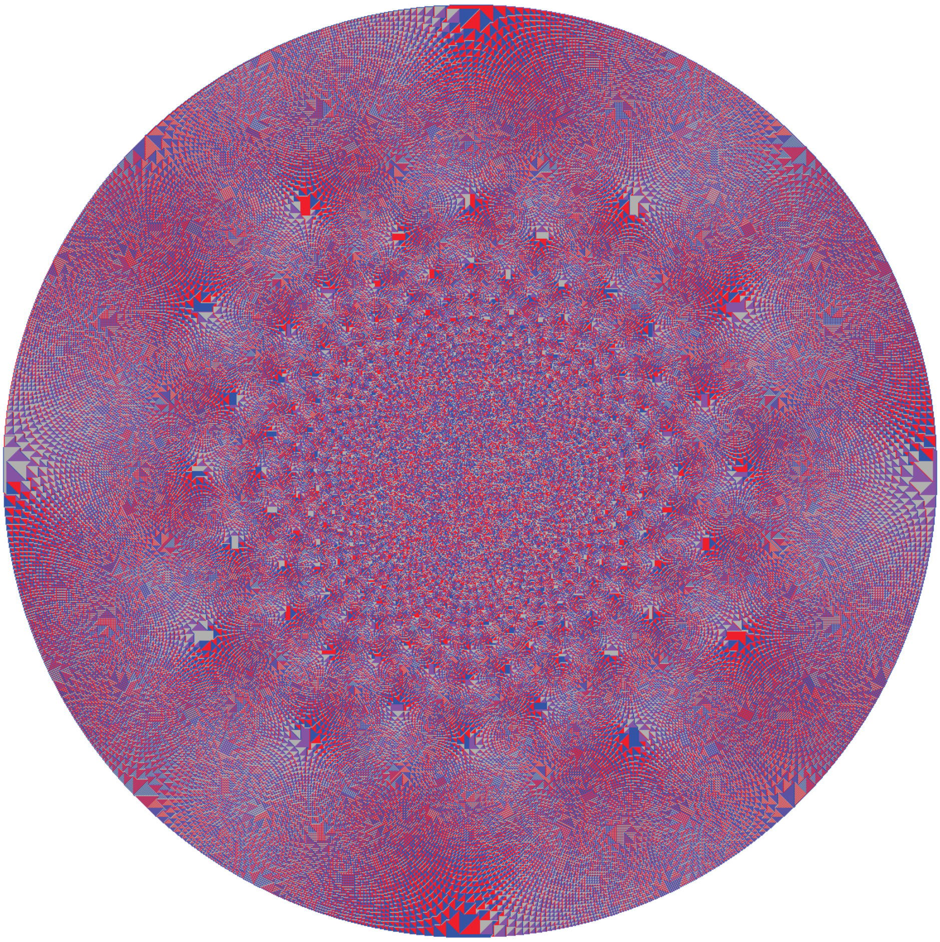

Figure 1: Rotor-Router shape created by 1 million particles starting at the origin, the sites are colored by the direction of the rotors. [L]

4.2 Classical Sandpile

Another deterministic growth model is the Classical Sandpile or sometimes called the Abelian sandpile. It was introduced in [Dh]. The simplest version of the Classical Sandpile is described by considering an initial discrete mass configuration σ as in Definition 1.1.1. A site topples if $\sigma(x) \geq 2d$, adding one particle to every neighbouring site and emptying x. As before, two neighbouring sites are x and y such that $|x \sim y| = 1$. We continue the topplings until every site satisfies $\sigma(x) < 2d$. We then generalize this by saying that every empty site (x such that $\sigma(x) = 0$) starts with a "hole" of depth H, meaning that a site absorbs the first H particles it recieves and then the site functions as normal, toppling once for every

$2d$ particles recieved. That is, a site x with $\sigma(x) = 0$ absorbs $2d+H$ particles before toppling the first time. When it does so it sends out just $2d$ particles. After the initial toppling it topples once for every 2d particles it recieves as normal. We can also consider the classical sandpile with negative H. This is defined by considering the usual model but with $-H$ particles starting at each site with $\sigma(x) = 0$, that is, given σ we add $-H$ particles to every empty site. We have to be careful since if $-H \geq 2d - 1$ the topplings will not stop if we add just a single grain. One can then ask if there exists an asymptotic shape and what it may be. The existence of an asymptotic shape is not known but we have some bounds of the domain of occupied sites.

Theorem 4.2.1. *Let $H \geq 2 - 2d$. Denote the domain of visited sites of the Classical Sandpile with parameter H run with n particles starting at the origin by $S_{n,H}$. Let r be defined by $n = \omega_d r^d$, where ω_d is the volume of the unit ball in d dimensions. Then we have for $H \geq 1 - d$ and for any $\epsilon > 0$*

$$B_{c_1 r - c_2} \subset S_{n,H} \subset B_{c_1' r + c_2'},$$

where $c_1 = (2d - 1 + H)^{-1/d}$, $c_1' = (d - \epsilon + H)^{-1/d}$ and c_2 and c_2' are independent of n. When $H < 1 - d$ we do not have the full result, however we have the first inclusion $B_{c_1 r - c_2} \subset S_{n,H}$.

Proof. See [L]. □

The proof strategy is to observe that the Divisible Sandpile will behave similarly to the Classical Sandpile, enabling us to use results for the Divisible Sandpile. However, simulations seem to suggest a much more complicated behaviour of the asymptotic shape.

There is a conjecture that for $d = 2$, the asymptotic shape of $S_{n,H}$ is a regular polygon with $4H + 12$ sides. For $H = -2$ it is known that the asymptotic shape is a square; a proof by can be found in [FR].

5 Appendix

5.1 The discrete and the continuous Laplace operator

The continuous Laplace operator Δ is commonly formulated in cartesian coordinates in $\mathbb{R}^d$ for C^2-functions by the following expression:

$$\Delta f(x) = \sum_{i=1}^{d} \frac{\partial^2 f}{\partial x_i^2},$$

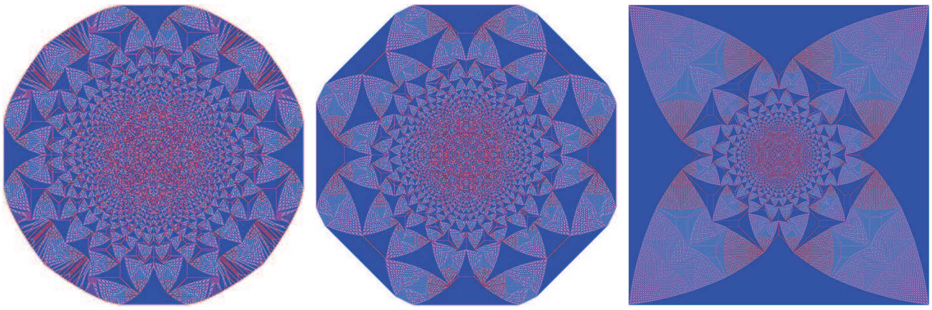

Figure 2: The picture shows $S_{n,0}$, $S_{n,-1}$ and $S_{n,-2}$ respectively for $n = 250000$ where the color indicates the amount of particles. [L]

and the discrete Laplacian of a function f defined on a graph is defined by

$$\Delta f(x) = \frac{1}{2d} \sum_{y \sim x} f(y) - f(x),$$

where $y \sim x$ means that y and x are neighbours (i.e. $|x \sim y| = 1$). Note that the scaling by $2d$ in the discrete case is not standard, but it is used here since it makes some correspondences more natural. The similarity of these operators is that they in some sense measure how the local average compares to the function value of that point. We can now define the notion of harmonic function:

Definition 5.1.1. *A function $f(x)$ is called harmonic at x if $\Delta f(x) = 0$, superharmonic if $\Delta f(x) \leq 0$ and subharmonic if $\Delta f(x) \geq 0$.*

In the continuous case we will extend the definitions to functions that are not in C^2.

Definition 5.1.2. *Define the mean value of a function over a ball of radius r as*

$$A_r f(x) := \frac{1}{\omega_d r^d} \int_{B(x,r)} f(y) dy.$$

Then if f is lower-semicontinuous and for any ball $A_r f(x) \leq f(x)$ the function is superharmonic and likewise if f is upper-semicontinuous and for any ball $A_r f(x) \geq f(x)$ then f is subharmonic. If a function is both super- and subharmonic it is called harmonic.

This generalization uses the mean-value property of harmonic functions.

5.2 Discrete maximum-minimum principle

Theorem 5.2.1. *Let f be a superharmonic (subharmonic) function defined in a subset $\Omega \subseteq \mathbb{Z}^d$, then f has no local minima (maxima) at interior points of Ω or it is constant in that connected component. By a local minimum we mean a point x so that $f(x) \leq f(y)$ for all neighbouring points y.*

Proof. We will prove the theorem for a superharmonic function, the subharmonic case is proven in the same way. Suppose f is superharmonic at x (for this to make sense it must be defined at all neighbouring points) and x is a local minimum for f. By the superharmonicity

$$\frac{1}{2d} \sum_{y \sim x} f(y) - f(x) \leq 0.$$

It follows that either all neighbouring points are equal or that at least one of the neighbouring points have less value. If all neighbouring points are equal we pick another neighbouring point and proceed by induction. If it has neighbouring points of less value it contradicts the assumption of local minimum. □

Corollary 5.2.1. *If f is superharmonic (subharmonic), defined on a finite set the minimum (maximum) is obtained on the boundary.*

5.3 Probabilistic concepts

5.3.1 Expected value

In this text we only use the expected value of functions on a probability space with at most countably many values. The definition in the continuous case is similar, substituting sums with integrals where needed.

Definition 5.3.1. *Let f be a real valued function on a probability space with at most countably many possible events $\{A_i\}_i$ with associated probability measure P. Then, if the following series is absolutely convergent, we define the expected value (expectation operator) as*

$$\mathbb{E}[f] = \sum_{i=1}^{\infty} f(A_i)P(A_i).$$

5.4 Stochastic processes

Definition 5.4.1. *A stochastic process with state space X is a collection of X-valued random variables indexed by a time-set T. In this work the time-set can be* *taken to be either* $\mathbb{N}$ *or* $\mathbb{R}$. *That is, a stochastic process as a collection of random variables*

$$\{F_t | t \in T\}.$$

A process where $T = \mathbb{N}$ *is called a discrete time process and if* $T = \mathbb{R}$ *it is called a continuous time process.*

5.4.1 Simple random walk

Definition 5.4.2. *A d-dimensional simple random walk is a* $\mathbb{Z}^d$*-valued discrete time stochastic process. We define it recursively. Choose a starting vertex* $x \in \mathbb{Z}^d$ *and set* $F_0 = x$ *with probability* 1. *Then define* F_t *so that the probability that* F_t *is at a site adjacent to* F_{t-1} *is* $\frac{1}{2d}$. *This is interpreted as a particle moving in* $\mathbb{Z}^d$ *so that it starts at* x *and then moves to an adjacent site at each time step. Every adjacent site is equally probable.*

5.4.2 Poisson process

Definition 5.4.3. *The Poisson process is an* $\mathbb{N}$*-valued continuous time process. It is defined to have the following properties, which defines it uniquely. First we demand that* $F_0 = 0$. *Secondly it should be nondecreasing and have independent increments, that is the behavior in disjoint time intervals are independent. Furthermore the probability distribution of increments in a given interval is only dependent of the length of the interval. Lastly, increments are always one at a time.*

The expected number of increments in an interval of unit length is called the intensity of the process.

Results and discussion about the Poisson process can be found for example in [Fe].

5.5 Green's function

5.5.1 Continuous Green's function

Definition 5.5.1. *For* $d = 2$ *we define*

$$g(x, y) = \frac{2}{\pi} \log |x - y|.$$

For $d \geq 3$ *we define*

$$g(x, y) = \frac{2}{(d-2)\omega_d} |x - y|^{2-d},$$

where ω_d *is the volume of d-dimensional unit ball.*

Note that this definitions differs by a factor $2d$ from the usual harmonic potential. This is chosen because it is convenient when working with the discrete Laplacian and random walks.

5.5.2 Discrete Green's function

We will construct the discrete version of Green's function using a random walk. Note however that this function is useful just for its properties as well and therefore can be used even when the original problem does not involve random walks. The relation with the continuous version can for example be seen as follows, if we translate this definition to the continuous case (i.e. Brownian motion) the corresponding function turns out to be that of Definition 5.5.1. For $d \geq 3$ a simple random walk is transient (see for example Theorem 4.1.1 of [LL]), meaning that the expected number of returns to a single point is finite. Thus we can define

$$g_1(x, y) = \mathbb{E}(\sum_{n=1}^{\infty} 1_{\{X_x(n)=y\}}),$$

where $\mathbb{E}$ is the expectation operator and $1_{\{X_x(n)=y\}}$ is a function that is 1 if the simple random walk $X_x(t)$ started at x is at site y after exactly n steps and 0 otherwise. We note that $g_1(x, y) = g_1(y, x) = g_1(x - y, 0)$ which should also be clear from the intuition about random walks. What we want to prove now is that

$$\Delta g_1(0, x) = -\delta_{0x},$$

where δ is Kronecker's delta.

Theorem 5.5.1. $\Delta g_1(0, x) = -\delta_{0x}$ *where* g_1 *is the discrete Green's function.*

Proof. We can write

$$g_1(0, x) = \mathbb{E}(\sum_{n=1}^{\infty} 1_{\{X_0(n)=x\}}) = 1_{\{x=0\}} + \sum_{y} p(y, x) g_1(y, 0)$$

where $p(x, y)$ is $\frac{1}{2d}$ if $x \sim y$ and 0 otherwise or equivalently the probability that a random walk jumps from y to x in one step. This identity is obtained by observing that for every site the random walk visits it arrived from an adjacent site or that it started there. Now applying Δ to both sides we see that $\Delta g_1(0, x) = -\delta_{0x}$ since the second term vanishes.

□

In the case $d = 2$ the expected number of visits of a simple random walk at a single site is unbounded (again see Theorem 4.1.1 of [LL]). We will therefore have to construct another function with similar properties in an alternative way. This function is called the *potential kernel* (note however that some authors call the negative of our function the potential kernel). We will use the same notation as for the discrete Green's function. It will be clear from the dimension number d which function we mean. So, for $d = 2$, let

$$g_1(x, y) = \sum_{n=0}^{\infty} (\mathbb{P}(X_x(n) = y) - \mathbb{P}(X_x(n) = x)) = \lim_{n\to\infty} [g_1^n(x, y) - g_1^n(x, x)]$$

where $g_1^n(x, y)$ is the expected number of visits at y in the n first steps of a simple random walk started at x. We have

$$|\mathbb{P}(X_x(n) = y) - \mathbb{P}(X_x(n) = x)| \le c|x - y|n^{-3/2} \tag{2}$$

making the sum $\sum_{n=1}^{\infty} \mathbb{P}(X_x(n) = y) - \mathbb{P}(X_x(n) = x)$ converge absolutely. However we leave out the proof.

Note also that

$$\Delta_y(\mathbb{P}(X_x(n) = y)) = \mathbb{P}(X_x(n + 1) = y) - \mathbb{P}(X_x(n) = y),$$

which is the key to proving that the potential kernel satisfies the same identity as Green's function. We want to prove that $\Delta g_1(0, x) = -\delta_{0x}$ for the potential kernel as well.

Theorem 5.5.2. $\Delta g_1(0, x) = -\delta_{0x}$, *where* g_1 *is the potential kernel.*

Proof.

$$\begin{aligned}
\Delta g_1(0, x) &= \Delta \sum_{n=0}^{\infty} (\mathbb{P}(X_0(n) = x) - \mathbb{P}(X_0(n) = 0)) \\
&= \sum_{n=0}^{\infty} \Delta(\mathbb{P}(X_0(n) = x) - \mathbb{P}(X_0(n) = 0)) \\
&= \lim_{N\to\infty} \sum_{n=0}^{N} \Delta(\mathbb{P}(X_0(n) = x) - \mathbb{P}(X_0(n) = 0)) \\
&= \lim_{N\to\infty} \sum_{n=0}^{N} (\mathbb{P}(X_0(n + 1) = x) - \mathbb{P}(X_0(n) = x)) \\
&= \lim_{N\to\infty} (\mathbb{P}(X_0(n + 1) = x) - \mathbb{P}(X_0(0) = x)) \\
&= -\mathbb{P}(X_0(0) = x) = -\delta_{0x},
\end{aligned}$$

where we have used the estimate (2) in the second equality. □

Next we are going to present a couple of asymptotic estimates for g_1 both for the case when $d = 2$ and when $d \geq 3$. These are standard estimates and can for example be found in [LL, Th. 4,3,1 and 4,4,4].

Theorem 5.5.3. *If $g_1(x, y)$ is the discrete Green's function in dimension $d \geq 3$, then*

$$g_1(x, y) = \frac{2}{(d-2)\omega_d}|x - y|^{2-d} + O(|x - y|^{-d}),$$

where ω_d is the volume of the unit ball in $\mathbb{R}^d$.

Theorem 5.5.4. *If $g_1(x, y)$ is the potential kernel in dimension $d = 2$, then*

$$g_1(x, y) = -\frac{2}{\pi} \log |x - y| + \frac{2\gamma + \log(8)}{\pi} + O(|x - y|^{-2}),$$

where $\gamma = \lim_{n\to\infty} \sum_{k=1}^{n}(1/k - \log(n)) = 0.577...$ is Euler's Constant.

The interesting thing to note is that g_1 and g are very close. This is crucial for our whole approach.

5.5.3 Green's function for a random walk in a ball

Definition 5.5.2. *Let $X_y(t)$ be a random walk started at y that is killed upon leaving the ball $B(0, n)$. Then let*

$$g_{B(0,n)}(y, z) = \mathbb{E}[\sum_{t=0}^{\infty} 1_{\{X_y(t)=z\}}].$$

Theorem 5.5.5. *We have the following estimates*

$$g_{B(0,n)}(0, z) = \frac{2}{\pi} \log \frac{n}{\max(|z|, 1)} + o(\frac{1}{\max(|z|, 1)}) + O(\frac{1}{n}),$$

for $d = 2$,

$$g_{B(0,n)}(0, z) = \frac{2}{(d-2)\omega_d}((\max(|z|, 1))^{2-d} - n^{2-d}) + O((\max(|z|, 1))^{1-d}),$$

for $d \geq 3$ and

$$g_{B(0,\epsilon n)}(0, 0) \leq g_{B(0,n)}(z, z) \leq g_{B(0,2n)}(0, 0),$$

for all d and $z \in B(0, n(1 - \epsilon))$.

Furthermore for $z \in B(0, n)$, we have

$$n^2 - (\max(|z|, 1))^2 \leq \mathbb{P}(\tau_n) \leq (n + 1)^2 - (\max(|z|, 1))^2,$$

where τ_n is the time where the random walk leaves $B(0,n)$.

Lastly, for n sufficiently large and $z \in B(0, n(1-\epsilon))$ we have

$$\sum_{y \in B(0,n)} g_{B(0,n)}(y,z) \leq \omega_d n^d g_{B(0,n)}(0,z).$$

Proof. See for example [LBG]. □

5.6 Borel-Cantelli

Borel-Cantelli's lemma is a statement in probability theory that is very useful when one wants to prove that something happens with probabilty 0.

Theorem 5.6.1. *Let $\{E_n\}_{n=1}^{\infty}$ be a sequence of events in some probability space. Then*

$$\sum_{n=1}^{\infty} \mathbb{P}(E_n) < \infty \Rightarrow \mathbb{P}(\limsup_n E_n) = 0.$$

Proof. See for example [Fe]. □

5.7 Sums of independent indicator variables

Definition 5.7.1. *A function defined defined on a probability space taking the values 0 and 1 is called a random indicator function.*

If we form a sum of independent indicator functions we would expect that this is not too far away from the sum of the expectations. The next theorem tells us exactly that.

Theorem 5.7.1. *Let S be a finite sum of independent indicator random functions with mean μ. Then for any $0 < \gamma < 1/2$ and for large enough μ one has*

$$\mathbb{P}(|S-\mu| \geq \mu^{1/2+\gamma}) \leq 2\exp(\frac{-\mu^{2\gamma}}{4}).$$

Proof. See for example [LBG]. □

5.8 Notations

$\mathbb{Z}^d$	The d-dimensional integers
T_i	A toppling function at i
δ_{ij}	Kronecker's delta function
σ	An initial mass configuration
ν	A final mass configuration
u	The odomenter function
Δ	Laplace operator, discrete or continuous
γ	An obstacle given by some initial configuration
s	A solution to an obstacle problem
$\mathbb{R}^d$	d-dimensional Euclidean space
$B(x,r)$	A ball of radius r around the point x or in Section 3 the lattice ball
A^ϵ	Outer ϵ-neighbourhood
A_ϵ	Inner ϵ-neighbourhood
$\delta_n\mathbb{Z}^d$	d-dimensional integers with grid size δ_n
$\lfloor a \rfloor$	The closest integer to a (rounding down)
$x^\square$	The set $x + [-\delta_n/2, \delta_n/2]^d$
$\mathbb{Z}_+$	{0,1,2,...}
$\bar{\Omega}$	Closure of the set Ω
Ω°	Interior of the set Ω
Ω^c	Complement of the set Ω
∇	Gradient operator
δ_{x_i}	Dirac delta function at x_i
ω_d	Volume of the unit ball in $\mathbb{R}^d$
$1_{\{E\}}$	Function with value 1 if E is true and 0 if it is false
$\mathbb{P}$	Probability operator
$\mathbb{E}$	Expectation operator
$X^i(t)$	The state of the i'th random walk after t steps
C^2	The class of twice continuosly differentiable functions
$g(x,y)$	Green's function
$g_1(x,y)$	Discrete Green's function or potential kernel
$g_{B(0,n)}$	Discrete Green's function for a ball of radius n

6 Acknowledgements

I would like to thank my supervisor Henrik Shahgholian for his support and for introducing me to this fascinating topic. I would also like to thank my dad for reading several drafts and for his valuable comments.

References

[C] M. Cook, *The tent metaphor* `http://paradise.caltech.edu/~cook/Warehouse/ForPropp/`

[Dh] D. Dhar *Self-Organized Critical State of Sandpile Automaton Models* , Phys. Rev. Lett. 64(14) (1990) , 1613-1616.

[DF] P. Diaconis and W. Fulton, *A growth model, a game, an algebra, Lagrange inversion, and characteristic classes*, Rend. Sem. Mat. Univ. Pol. Torino 49 (1991) no. 1, 95119.

[Do] J. Doob *Classical potential theory and its probabilistic counterpart* Reprint of the 1984 edition. Classics in Mathematics. Springer-Verlag, Berlin, 2001. xxvi+846 pp. ISBN: 3-540-41206-9.

[Fe] W. Feller, *An Introduction to Probability Theory and Its Applications*, John Wiley & Sons.

[FR] A. Fey and F.Redig, *Limiting shapes for deterministic centrally seeded growth models*, `http://arxiv.org/abs/math.PR/0702450`.

[F] A. Friedman, *Variational principles and free-boundary problems* A Wiley-Interscience Publication. Pure and Applied Mathematics. John Wiley & Sons, Inc., New York, 1982. ix+710 pp. ISBN: 0-471-86849-3.

[GQ] J. Gravner and J. Quastel, *Internal DLA and the Stefan problem*, *Ann. Probab.* **28**, no. 4 (2000), 1528–1562.

[LBG] G. Lawler, M. Bramson and D. Griffeath, *Internal diffusion limited aggregation*, *Ann. Probab.* **20**, no. 4 (1992), 2117–2140.

[LL] G. Lawler and V. Limic *Random walk: A modern introduction*, Cambridge Studies in Advanced Mathematics, 123. Cambridge University Press, Cambridge, 2010. xii+364 pp. ISBN: 978-0-521-51918-2.

[L] L. Levine *Limit theorems for internal aggregation models*, Thesis (Ph.D.)University of California, Berkeley. 2007. 132 pp. ISBN: 978-0549-52990-3.

[PDDK] V. B. Priezzhev, D. Dhar, A. Dhar, and S. Krishnamurthy *Eulerian walkers as a model of self-organised criticality* , *Phys. Rev. Lett. 77* (1996) 507982.