Limit Theorems for Internal Aggregation Models

by

Lionel Timothy Levine

A.B. Harvard University, 2002

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Mathematics

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:

Professor Yuval Peres, Chair
Professor Lawrence C. Evans
Professor Elchanan Mossel

Fall 2007

The dissertation of Lionel Timothy Levine is approved:

_____

Chair   Date

_____

                                                                               Date

_____

                                                                               Date

University of California, Berkeley

Fall 2007

Limit Theorems for Internal Aggregation Models

Copyright 2007

by

Lionel Timothy Levine

Abstract

Limit Theorems for Internal Aggregation Models

by

Lionel Timothy Levine

Doctor of Philosophy in Mathematics

University of California, Berkeley

Professor Yuval Peres, Chair

We study the scaling limits of three different aggregation models on $\mathbb{Z}^d$: internal DLA, in which particles perform random walks until reaching an unoccupied site; the rotor-router model, in which particles perform deterministic analogues of random walks; and the divisible sandpile, in which each site distributes its excess mass equally among its neighbors. As the lattice spacing tends to zero, all three models are found to have the same scaling limit, which we describe as the solution to a certain PDE free boundary problem in $\mathbb{R}^d$. In particular, internal DLA has a deterministic scaling limit. We find that the scaling limits are quadrature domains, which have arisen independently in many fields such as potential theory and fluid dynamics. Our results apply both to the case of multiple point sources and to the Diaconis-Fulton smash sum of domains.

In the special case when all particles start at a single site, we show that the scaling limit is a Euclidean ball in $\mathbb{R}^d$ and give quantitative bounds on the rate of convergence to a ball. For the divisible sandpile, the error in the radius is bounded by a constant independent of the total starting mass. For the rotor-router model in $\mathbb{Z}^d$, the inner error grows at most logarithmically in the radius r, while the outer error is at most order $r^{1-1/d} \log r$. We also improve on the previously best known bounds of Le Borgne and Rossin in $\mathbb{Z}^2$ and Fey and Redig in higher dimensions for the shape of the classical abelian sandpile model.

Lastly, we study the sandpile group of a regular tree whose leaves are collapsed to a single sink vertex, and determine the decomposition of the full sandpile group as a product of cyclic groups. For the regular ternary tree of height n, for example, the sandpile

group is isomorphic to $(\mathbb{Z}_3)^{2^{n-3}} \oplus (\mathbb{Z}_7)^{2^{n-4}} \oplus \ldots \oplus \mathbb{Z}_{2^{n-1}-1} \oplus \mathbb{Z}_{2^n-1}$. We use this result to prove that rotor-router aggregation on the regular tree yields a perfect ball.

To my parents, Lance and Terri.

Acknowledgments

This work would not have been possible without the incredible support and insight of my advisor, Yuval Peres. Yuval taught me not only a lot of mathematics, but also the tools, techniques, instincts and heuristics essential to the working mathematician. Somehow, he even managed to teach me a little bit about life as well. My entire philosophy and approach, and my sense of what is important in mathematics, are colored by his ideas.

I would like to thank Jim Propp for introducing me to this beautiful area of mathematics, and for many inspiring conversations over the years. Thanks also to Scott Armstrong, Henry Cohn, Darren Crowdy, Craig Evans, Anne Fey, Chris Hillar, Wilfried Huss, Itamar Landau, Karola Mészáros, Chandra Nair, David Pokorny, Oded Schramm, Scott Sheffield, Misha Sodin, Kate Stange, Richard Stanley, Parran Vanniasegaram, Grace Wang, and David Wilson for many helpful conversations. Richard Liang taught me how to write image files using C, so that I could write programs to generate many of the figures. In addition, Itamar Landau and Yelena Shvets helped create several of the figures.

I also thank the NSF for supporting me with a Graduate Research Fellowship during much of the time period when this work was carried out.

Finally, I would like to thank my family and friends for supporting me and believing in me through the best and worst of times. Your love and support mean more to me than I can possibly express in words.

Contents

Index of Notation

Chapter 1

Introduction

1.1 Three Models with the Same Scaling Limit

Given finite sets $A, B \subset \mathbb{Z}^d$, Diaconis and Fulton [15] defined the *smash sum* $A \oplus B$ as a certain random set whose cardinality is the sum of the cardinalities of A and B. Write $A \cap B = \{x_1, \ldots, x_k\}$. To construct the smash sum, begin with the union $C_0 = A \cup B$ and for each $j = 1, \ldots, k$ let

$$C_j = C_{j-1} \cup \{y_j\}$$

where y_j is the endpoint of a simple random walk started at x_j and stopped on exiting C_{j-1}. Then define $A \oplus B = C_k$. The key observation of [15] is that the law of $A \oplus B$ does not depend on the ordering of the points x_j. The sum of two squares in $\mathbb{Z}^2$ overlapping in a smaller square is pictured in Figure 1.1.

In Theorem 1.1.3, below, we prove that as the lattice spacing goes to zero, the smash sum $A \oplus B$ has a deterministic scaling limit in $\mathbb{R}^d$. Before stating our main results, we describe some related models and describe our technique for identifying their common scaling limit, which comes from the theory of free boundary problems in PDE.

The Diaconis-Fulton smash sum generalizes the model of *internal diffusion-limited aggregation* ("internal DLA") studied in [29], and in fact was part of the original motivation for that paper. In classical internal DLA, we start with n particles at the origin $o \in \mathbb{Z}^d$ and let each perform a simple random walk until it reaches an unoccupied site. The resulting random set of n occupied sites in $\mathbb{Z}^d$ can be described as the n-fold smash sum of $\{o\}$ with itself. We will use the term internal DLA to refer to particles which perform simple
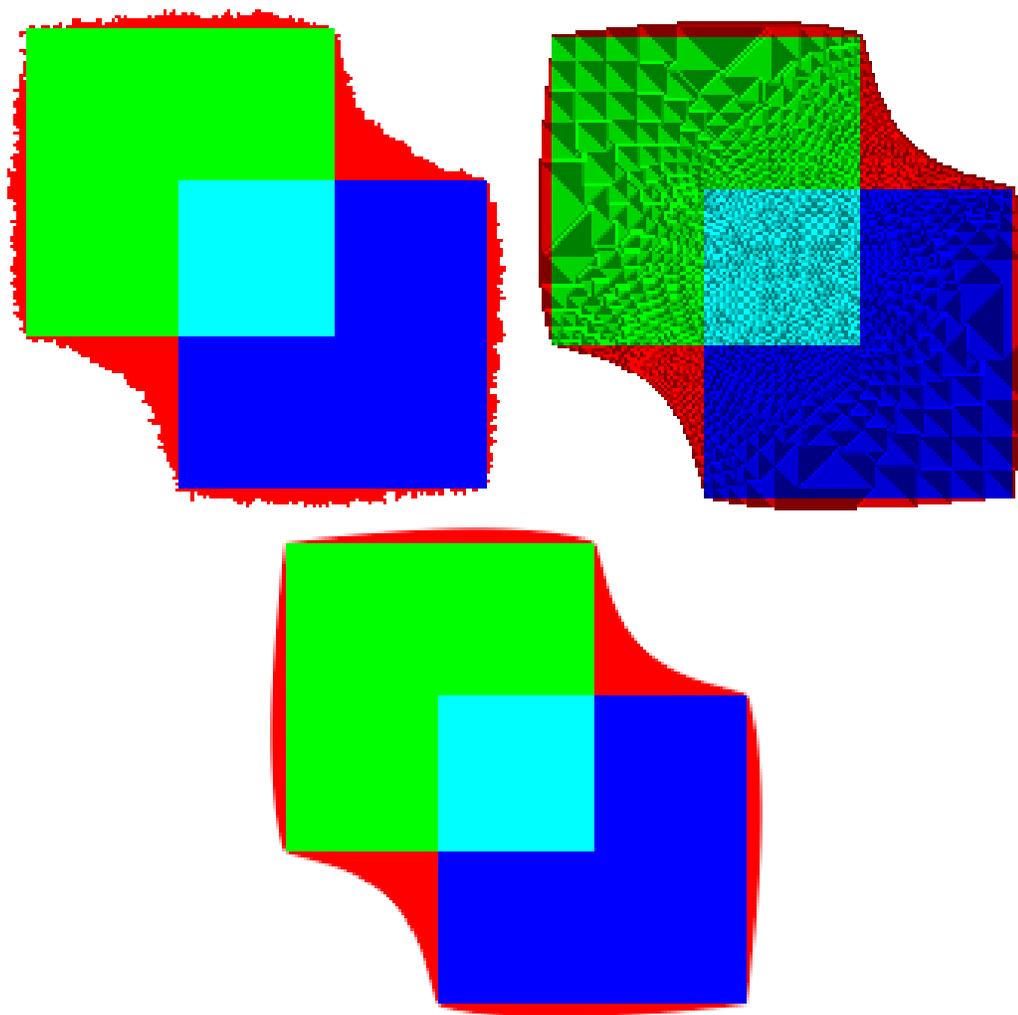
Figure 1.1: Smash sum of two squares in $\mathbb{Z}^2$ overlapping in a smaller square, for internal DLA (top left), the rotor-router model (top right), and the divisible sandpile.

random walks in $\mathbb{Z}^d$ until reaching an unoccupied site, starting from an arbitrary initial configuration. In this broader sense of the term, both the Diaconis-Fulton sum and the model studied in [29] are particular cases of internal DLA.

In defining the smash sum $A \oplus B$, various alternatives to random walk are possible. *Rotor-router walk* is a deterministic analogue of random walk, first studied by Priezzhev et al. [38] under the name "Eulerian walkers." At each site in $\mathbb{Z}^2$ is a *rotor* pointing north, south, east or west. A particle performs a nearest-neighbor walk on the lattice according to the following rule: during each time step, the rotor at the particle's current location is

rotated clockwise by 90 degrees, and the particle takes a step in the direction of the newly rotated rotor. In higher dimensions, the model can be defined analogously by repeatedly cycling the rotors through an ordering of the $2d$ cardinal directions in $\mathbb{Z}^d$. The sum of two squares in $\mathbb{Z}^2$ using rotor-router walk is pictured in Figure 1.1; all rotors began pointing west. The shading in the figure indicates the final rotor directions, with four different shades corresponding to the four possible directions.

The *divisible sandpile* model uses continuous amounts of mass in place of discrete particles. A lattice site is *full* if it has mass at least 1. Any full site can *topple* by keeping mass 1 for itself and distributing the excess mass equally among its neighbors. At each time step, we choose a full site and topple it. As time goes to infinity, provided each full site is eventually toppled, the mass approaches a limiting distribution in which each site has mass ≤ 1. Note that individual topplings do not commute. However, the divisible sandpile is "abelian" in the sense that any sequence of topplings produces the same limiting mass distribution; this is proved in Lemma 2.2.1. Figure 1.1 shows the limiting domain of occupied sites resulting from starting mass 1 on each of two squares in $\mathbb{Z}^2$, and mass 2 on the smaller square where they intersect.

Figure 1.1 raises a few natural questions: as the underlying lattice spacing becomes finer and finer, will the smash sum $A \oplus B$ tend to some limiting shape in $\mathbb{R}^d$, and if so, what is this shape? Will it be the same limiting shape for all three models? To see how we might identify the limiting shape, consider the divisible sandpile *odometer function*

$$u(x) = \text{total mass emitted from } x. \tag{1.1}$$

Since each neighbor $y \sim x$ emits an equal amount of mass to each of its $2d$ neighbors, the total mass received by x from its neighbors is $\frac{1}{2d} \sum_{y \sim x} u(y)$, hence

$$\Delta u(x) = \nu(x) - \sigma(x) \tag{1.2}$$

where $\sigma(x)$ and $\nu(x)$ are the initial and final amounts of mass at x, respectively. Here Δ is the discrete Laplacian in $\mathbb{Z}^d$, defined by

$$\Delta u(x) = \frac{1}{2d} \sum_{y \sim x} u(y) - u(x). \tag{1.3}$$

Equation (1.2) suggests the following approach to finding the limiting shape. We first construct a function on $\mathbb{Z}^d$ whose Laplacian is $\sigma - 1$; an example is the function

$$\gamma(x) = -|x|^2 - \sum_{y \in \mathbb{Z}^d} g_1(x, y)\sigma(y) \tag{1.4}$$

where in dimension $d \geq 3$ the Green's function $g_1(x, y)$ is the expected number of times a simple random walk started at x visits y (in dimension $d = 2$ we use the recurrent potential kernel in place of the Green's function). Here $|x|$ denotes the Euclidean norm $(x_1^2 + \ldots + x_d^2)^{1/2}$. By (1.2), since $\nu \leq 1$ the sum $u + \gamma$ is a superharmonic function on $\mathbb{Z}^d$; that is, $\Delta(u + \gamma) \leq 0$. Moreover if $f \geq \gamma$ is any superharmonic function lying above γ, then $f - \gamma - u$ is superharmonic on the domain $D = \{x \in \mathbb{Z}^d | \nu(x) = 1\}$ of fully occupied sites, and nonnegative outside D, hence nonnegative everywhere. Thus we have proved the following lemma.

Lemma 1.1.1. *Let σ be a nonnegative function on $\mathbb{Z}^d$ with finite support. Then the odometer function (1.1) for the divisible sandpile started with mass $\sigma(x)$ at each site x is given by*

$$u = s - \gamma$$

where γ is given by (1.4), and

$$s(x) = \inf\{f(x) | f \text{ is superharmonic on } \mathbb{Z}^d \text{ and } f \geq \gamma\}$$

is the least superharmonic majorant of γ.

Lemma 1.1.1 allows us to formulate the problem in a way which translates naturally to the continuum. Given a function σ on $\mathbb{R}^d$ representing the initial mass density, by analogy with (1.4) we define the *obstacle*

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

where $g(x, y)$ is the Green's function on $\mathbb{R}^d$ proportional to $|x - y|^{2-d}$ in dimensions $d \geq 3$ and to $-\log|x - y|$ in dimension two. We then let

$$s(x) = \inf\{f(x) | f \text{ is continuous, superharmonic and } f \geq \gamma\}.$$

The *odometer function* for σ is then given by $u = s - \gamma$, and the final domain of occupied sites is given by

$$D = \{x \in \mathbb{R}^d | s(x) > \gamma(x)\}. \tag{1.5}$$

This domain D is called the *noncoincidence set for the obstacle problem* with obstacle γ; for an in-depth discussion of the obstacle problem, see [19].
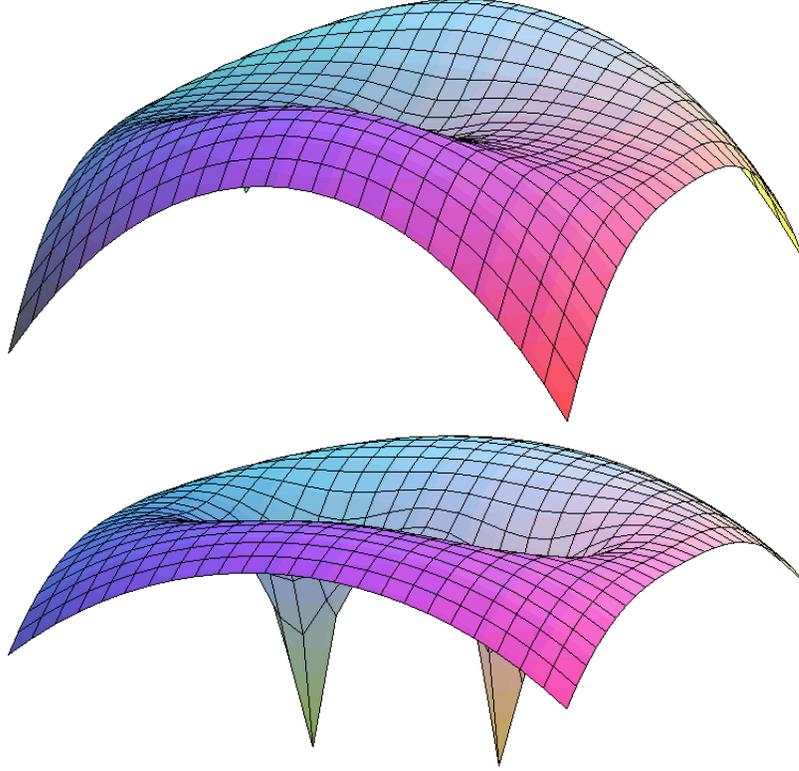
Figure 1.2: The obstacles γ corresponding to starting mass 1 on each of two overlapping disks (top) and mass 100 on each of two nonoverlapping disks.

If A, B are bounded open sets in $\mathbb{R}^d$, we define the *smash sum* of A and B as

$$A \oplus B = A \cup B \cup D \tag{1.6}$$

where D is given by (1.5) with $\sigma = 1_A + 1_B$. In the two-dimensional setting, an alternative definition of the smash sum in terms of quadrature identities is mentioned in [23].

In this thesis we prove, among other things, that if any of our three aggregation models is run on finer and finer lattices with initial mass densities converging in an appropriate sense to σ, the resulting domains of occupied sites will converge in an appropriate sense to the domain D given by (1.5). We will always work in dimension $d \geq 2$; for a discussion of the rotor-router model in one dimension, see [32].

Let us define the appropriate notion of convergence of domains, which amounts essentially to convergence in the Hausdorff metric. Fix a sequence $\delta_n \downarrow 0$ representing the

lattice spacing. Given domains $A_n \subset \delta_n \mathbb{Z}^d$ and $D \subset \mathbb{R}^d$, write $A_n \to D$ if for any $\epsilon > 0$

$$D_\epsilon \cap \delta_n \mathbb{Z}^d \subset A_n \subset D^\epsilon \tag{1.7}$$

for all sufficiently large n. Here

$$D_\epsilon = \{x \in D \,|\, B(x, \epsilon) \subset D\} \tag{1.8}$$

and

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

are the inner and outer ϵ-neighborhoods of D. For $x \in \delta_n \mathbb{Z}^d$ we write $x^\square = \left[x + \frac{\delta_n}{2}, x - \frac{\delta_n}{2}\right]^d$. For $t \in \mathbb{R}$ write $\lfloor t \rceil$ for the closest integer to t.

Throughout this thesis, to avoid trivialities we work in dimension $d \geq 2$. Our first main result is the following.

Theorem 1.1.2. *Let $\Omega \subset \mathbb{R}^d$ be a bounded open set, and let $\sigma : \mathbb{R}^d \to \mathbb{Z}_{\geq 0}$ be a bounded function which is continuous almost everywhere, satisfying $\{\sigma \geq 1\} = \bar{\Omega}$. Let D_n, R_n, I_n be the domains of occupied sites formed from the divisible sandpile, rotor-router model, and internal DLA, respectively, in the lattice $\delta_n \mathbb{Z}^d$ started from source density*

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

Then as $n \to \infty$

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

and if $\delta_n \leq 1/\log n$, then with probability one

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

where D is given by (1.5), and the convergence is in the sense of (1.7).

Remark. When forming the rotor-router domains R_n, the initial rotors in each lattice $\delta_n \mathbb{Z}^d$ may be chosen arbitrarily.

We prove a somewhat more general form of Theorem 1.1.2 which allows for some flexibility in how the discrete density σ_n is constructed from σ. In particular, taking $\sigma = 1_{\bar{A}} + 1_{\bar{B}}$ we obtain the following theorem, which explains the similarity of the three smash sums pictured in Figure 1.1.

Theorem 1.1.3. *Let $A, B \subset \mathbb{R}^d$ be bounded open sets whose boundaries have measure zero. Let D_n, R_n, I_n be the smash sum of $A \cap \delta_n \mathbb{Z}^d$ and $B \cap \delta_n \mathbb{Z}^d$, formed using divisible sandpile, rotor-router and internal DLA dynamics, respectively. Then as $n \to \infty$*

$$D_n, R_n \to A \oplus B;$$

and if $\delta_n \leq 1/\log n$, then with probability one

$$I_n \to A \oplus B$$

where $A \oplus B$ is given by (1.6), and the convergence is in the sense of (1.7).

For the divisible sandpile, Theorem 1.1.2 can be generalized by dropping the requirement that σ be integer valued; see Theorem 3.2.7 for the precise statement. Taking σ real-valued is more problematic in the case of the rotor-router model and internal DLA, since these models work with discrete particles. Still, one might wonder if, for example, given a domain $A \subset \mathbb{R}^d$, starting each even site in $A \cap \delta_n \mathbb{Z}^d$ with one particle and each odd site with two particles, the resulting domains R_n, I_n would converge to the noncoincidence set D for density $\sigma = \frac{3}{2} 1_A$. This is in fact the case: if σ_n is a density on $\delta_n \mathbb{Z}^d$, as long as a certain "smoothing" of σ_n converges to σ, the rotor-router and internal DLA domains started from source density σ_n will converge to D. See Theorems 3.3.7 and 3.4.1 for the precise statements.

1.2 Single Point Sources

One interesting case not covered by Theorems 1.1.2 and 1.1.3 is the case of point sources. Lawler, Bramson and Griffeath [29] showed that the scaling limit of internal DLA in $\mathbb{Z}^d$ with a single point source of particles is a Euclidean ball. In chapter 2, we prove analogous results for rotor-router aggregation and the divisible sandpile, and give quantitative bounds on the rate of convergence to a ball. Let A_n be the domain of n sites in $\mathbb{Z}^d$ formed from rotor-router aggregation starting from a point source of n particles at the origin. Thus A_n is defined inductively by the rule

$$A_n = A_{n-1} \cup \{x_n\}$$

where x_n is the endpoint of a rotor-router walk started at the origin in $\mathbb{Z}^d$ and stopped on first exiting A_{n-1}. For example, in $\mathbb{Z}^2$, if all rotors initially point north, the sequence will
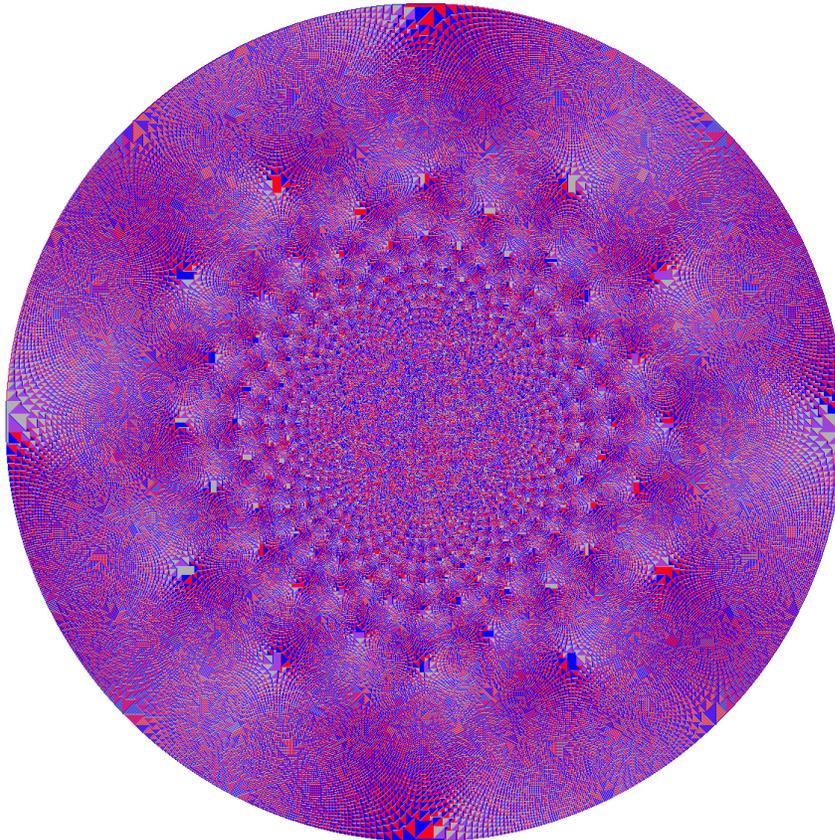
Figure 1.3: Rotor-router aggregate of one million particles in $\mathbb{Z}^2$. Each site is colored according to the direction of its rotor.

begin $A_1 = \{(0,0)\}$, $A_2 = \{(0,0), (1,0)\}$, $A_3 = \{(0,0), (1,0), (0,-1)\}$. The region $A_{1,000,000}$ is pictured in Figure 1.3.

Jim Propp observed from simulations in two dimensions that the regions A_n are extraordinarily close to circular, and asked why this was so [26, 39]. Despite the impressive empirical evidence for circularity, the best result known until now [33] says only that if A_n is rescaled to have unit volume, the volume of the symmetric difference of A_n with a ball of unit volume tends to zero as a power of n, as $n \uparrow \infty$. The main outline of the argument is summarized in [34]. Fey and Redig [18] also show that A_n contains a diamond. In particular, these results do not rule out the possibility of "holes" in A_n far from the boundary or of long tendrils extending far beyond the boundary of the ball, provided the volume of these features is negligible compared to n.

Our main result on the shape of rotor-router aggregation with a single point source

is the following, which rules out the possibility of holes far from the boundary or of long tendrils. For $r \geq 0$ let

$$B_r = \{x \in \mathbb{Z}^d \ : \ |x| < r\}.$$

Theorem 1.2.1. *Let A_n be the region formed by rotor-router aggregation in $\mathbb{Z}^d$ starting from n particles at the origin and any initial rotor state. There exist constants c, c' depending only on d, such that*

$$B_{r-c\log r} \subset A_n \subset B_{r(1+c'r^{-1/d}\log r)}$$

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

We remark that the same result holds when the rotors evolve according to stacks of bounded discrepancy; see the remark following Lemma 2.4.1.

By way of comparison with Theorem 1.2.1, if I_n is the internal DLA region formed from n particles started at the origin, the best known bounds [30] are (up to logarithmic factors)

$$B_{r-r^{1/3}} \subset I_n \subset B_{r+r^{1/3}}$$

for all sufficiently large n, with probability one.

Our next result treats the divisible sandpile with all mass initially concentrated at a point source. The resulting domain of fully occupied sites is extremely close to a ball; in fact, the error in the radius is bounded independent of the total mass.

Theorem 1.2.2. *For $m \geq 0$ let $D_m \subset \mathbb{Z}^d$ be the domain of fully occupied sites for the divisible sandpile formed from a pile of mass m at the origin. There exist constants c, c' depending only on d, such that*

$$B_{r-c} \subset D_m \subset B_{r+c'},$$

where $r = (m/\omega_d)^{1/d}$ and ω_d is the volume of the unit ball in $\mathbb{R}^d$.

The divisible sandpile is similar to the "oil game" studied by Van den Heuvel [49]. In the terminology of [18], it also corresponds to the $h \to -\infty$ limit of the classical abelian sandpile (defined below), that is, the abelian sandpile started from the initial condition in which every site has a very deep "hole."
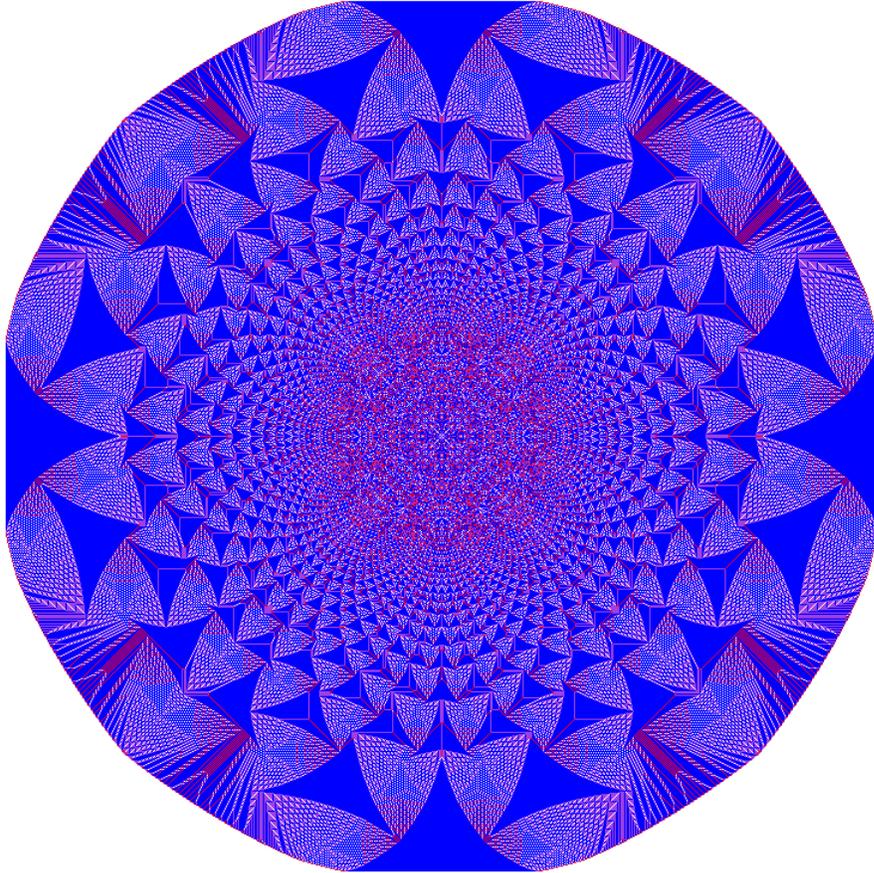
Figure 1.4: Classical abelian sandpile aggregate of one million particles in $\mathbb{Z}^2$. Colors represent the number of grains at each site.

In the classical *abelian sandpile model* [4], each site in $\mathbb{Z}^d$ has an integer number of grains of sand; if a site has at least $2d$ grains, it *topples*, sending one grain to each neighbor. If n grains of sand are started at the origin in $\mathbb{Z}^d$, write S_n for the set of sites that are visited during the toppling process; in particular, although a site may be empty in the final state, we include it in S_n if it was occupied at any time during the evolution to the final state.

Until now the best known constraints on the shape of S_n in two dimensions were due to Le Borgne and Rossin [31], who proved that

$$\{x \in \mathbb{Z}^2 \mid x_1 + x_2 \leq \sqrt{n/12} - 1\} \subset S_n \subset \{x \in \mathbb{Z}^2 \mid x_1, x_2 \leq \sqrt{n}/2\}.$$

Fey and Redig [18] proved analogous bounds in higher dimensions, and extended these bounds to arbitrary values of the height parameter h. This parameter is discussed in
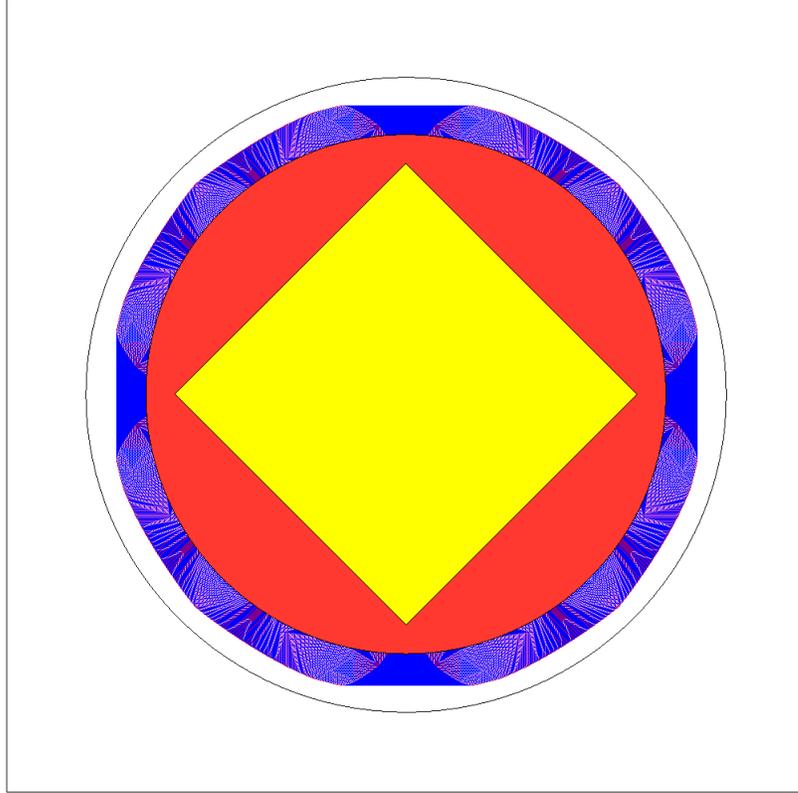
Figure 1.5: Known bounds on the shape of the classical abelian sandpile in $\mathbb{Z}^2$. The inner diamond and outer square are due to Le Borgne and Rossin [31]; the inner and outer circles are those in Theorem 1.2.3.

section 2.3.

The methods used to prove the near-perfect circularity of the divisible sandpile shape in Theorem 1.2.2 can be used to give constraints on the shape of the classical abelian sandpile, improving on the bounds of [18] and [31].

Theorem 1.2.3. *Let S_n be the set of sites that are visited by the classical abelian sandpile model in $\mathbb{Z}^d$, starting from n particles at the origin. Write $n = \omega_d r^d$. Then for any $\epsilon > 0$ we have*

$$B_{c_1 r - c_2} \subset S_n \subset B_{c_1' r + c_2'}$$

where

$$c_1 = (2d-1)^{-1/d}, \qquad c_1' = (d-\epsilon)^{-1/d}.$$

The constant c_2 depends only on d, while c_2' depends only on d and ϵ.
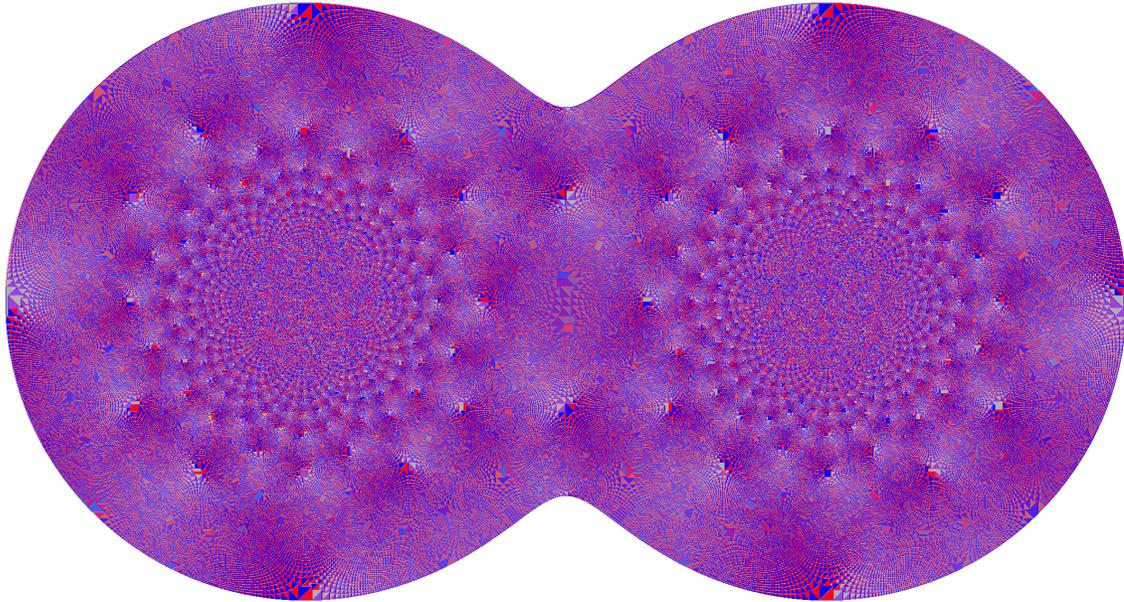
Figure 1.6: The rotor-router model in $\mathbb{Z}^2$ started from two point sources on the x-axis. The boundary of the limiting shape is an algebraic curve of degree 4; see equation (1.13).

Note that Theorem 1.2.3 does not settle the question of the asymptotic shape of S_n, and Figure 1.4 indicates that the limiting shape in two dimensions may be a polygon rather than a disk. Even the existence of an asymptotic shape is not known, however.

1.3 Multiple Point Sources

Using our results for single point sources together with the construction of the smash sum (1.6), we can understand the limiting shape of our aggregation models started from multiple point sources. The answer turns out to be a smash sum of balls centered at the sources. For $x \in \mathbb{R}^d$ write $x^{\cdots}$ for the closest lattice point in $\delta_n \mathbb{Z}^d$, breaking ties to the right. Our shape theorem for multiple point sources, which is deduced from Theorems 1.1.3, 1.2.1 and 1.2.2 along with the main result of [29], is the following.

Theorem 1.3.1. *Fix $x_1, \ldots, x_k \in \mathbb{R}^d$ and $\lambda_1, \ldots, \lambda_k > 0$. Let B_i be the ball of volume λ_i centered at x_i. Fix a sequence $\delta_n \downarrow 0$, and for $x \in \delta_n \mathbb{Z}^d$ let*

$$\sigma_n(x) = \left\lfloor \delta_n^{-d} \sum_{i=1}^{k} \lambda_i 1_{\{x = x_i^{\cdots}\}} \right\rfloor .$$

Let D_n, R_n, I_n be the domains of occupied sites in $\delta_n \mathbb{Z}^d$ formed from the divisible sandpile, rotor-router model, and internal DLA, respectively, started from source density σ_n. Then as $n \to \infty$

$$D_n, R_n \to B_1 \oplus \ldots \oplus B_k; \tag{1.9}$$

and if $\delta_n \leq 1/n$, then with probability one

$$I_n \to B_1 \oplus \ldots \oplus B_k$$

where $\oplus$ denotes the smash sum (1.6), and the convergence is in the sense of (1.7).

Implicit in equation (1.9) is the associativity of the smash sum operation, which is not readily apparent from the definition (1.6). For a proof of associativity, see Lemma 3.5.1. For related results in dimension two, see [41, Prop. 3.10] and [50, section 2.4].

We remark that a similar model of internal DLA with multiple point sources was studied by Gravner and Quastel [21], who also obtained a variational solution. In their model, instead of starting with a fixed number of particles, each source x_i emits particles according to a Poisson process. The shape theorems of [21] concern convergence in the sense of volume, which is a weaker form of convergence than (1.7).

1.4 Quadrature Domains

By analogy with the discrete case, we would expect that volumes add under the smash sum operation; that is,

$$\mathcal{L}(A \oplus B) = \mathcal{L}(A) + \mathcal{L}(B)$$

where $\mathcal{L}$ denotes Lebesgue measure in $\mathbb{R}^d$. Although this additivity is not immediately apparent from the definition (1.6), it holds for all bounded open $A, B \subset \mathbb{R}^d$ provided their boundaries have measure zero; see Corollary 3.1.14.

We can derive a more general class of identities known as *quadrature identities* involving integrals of harmonic functions over $A \oplus B$. Let us first consider the discrete case. If h is a superharmonic function on $\mathbb{Z}^d$, and σ is a mass configuration for the divisible sandpile (so each site $x \in \mathbb{Z}^d$ has mass $\sigma(x)$), the sum $\sum_{x \in \mathbb{Z}^d} h(x)\sigma(x)$ can only decrease when we perform a toppling. Thus

$$\sum_{x \in \mathbb{Z}^d} h(x)\nu(x) \leq \sum_{x \in \mathbb{Z}^d} h(x)\sigma(x), \tag{1.10}$$

where ν is the final mass configuration. We therefore expect the domain D given by (1.5) to satisfy the *quadrature inequality*

$$\int_D h(x)dx \leq \int_D h(x)\sigma(x)dx \tag{1.11}$$

for all integrable superharmonic functions h on D. For a proof under suitable smoothness assumptions on σ and h, see Proposition 3.1.11; see also [42].

A domain $D \subset \mathbb{R}^d$ satisfying an inequality of the form (1.11) is called a *quadrature domain* for σ. Such domains are widely studied in potential theory and have a variety of applications in fluid dynamics [9, 40]. For more on quadrature domains and their connection with the obstacle problem, see [1, 8, 24, 25, 42, 43]. Equation (1.10) can be regarded as a discrete analogue of a quadrature inequality; in this sense our aggregation models, in particular the divisible sandpile, produce discrete analogues of quadrature domains.

In Proposition 3.5.5, we show that the smash sum of balls $B_1 \oplus \ldots \oplus B_k$ arising in Theorem 1.3.1 obeys the classical quadrature identity

$$\int_{B_1 \oplus \ldots \oplus B_k} h(x)dx = \sum_{i=1}^{k} \lambda_i h(x_i) \tag{1.12}$$

for all harmonic functions h on $B_1 \oplus \ldots \oplus B_k$. This can be regarded as a generalization of the classical mean value property of harmonic functions, which corresponds to the case $k = 1$. Using results of Gustafsson [22] and Sakai [42] on quadrature domains in the plane, we can deduce the following theorem, which is proved in section 3.5.

Theorem 1.4.1. *Let $B_1, \ldots, B_k$ be disks in $\mathbb{R}^2$ with distinct centers. The boundary of the smash sum $B_1 \oplus \ldots \oplus B_k$ lies on an algebraic curve of degree $2k$. More precisely, there is a polynomial $P \in \mathbb{R}[x_1, x_2]$ of the form*

$$P(x_1, x_2) = \left(x_1^2 + x_2^2\right)^k + \text{lower order terms}$$

and there is a finite set of points $E \subset \mathbb{R}^2$, possibly empty, such that

$$\partial(B_1 \oplus \ldots \oplus B_k) = \{(x_1, x_2) \in \mathbb{R}^2 | P(x_1, x_2) = 0\} - E.$$

For example, if B_1 and B_2 are disks of equal radius $r > 1$ centered at $(1, 0)$ and $(-1, 0)$, then $\partial(B_1 \oplus B_2)$ is given by the quartic curve [44]

$$\left(x_1^2 + x_2^2\right)^2 - 2r^2\left(x_1^2 + x_2^2\right) - 2(x_1^2 - x_2^2) = 0. \tag{1.13}$$

This curve describes the shape of the rotor-router model with two point sources pictured in Figure 1.6.

1.5 Aggregation on Trees

Let T be the infinite d-regular tree. To define rotor-router walk on a tree, for each vertex of T choose a cyclic ordering of its d neighbors. Each vertex is assigned a rotor which points to one of the neighboring vertices, and a particle walks by first rotating the rotor at each site it comes to, then stepping in the direction of the newly rotated rotor. Fix a vertex $o \in T$ called the origin. Beginning with $A_1 = \{o\}$, define the rotor-router aggregation cluster A_n inductively by

$$A_n = A_{n-1} \cup \{x_n\}, \qquad n > 1$$

where $x_n \in T$ is the endpoint of a rotor-router walk started at o and stopped on first exiting A_{n-1}. We do not change the positions of the rotors when adding a new chip. Thus the sequence $(A_n)_{n \geq 1}$ depends only on the choice of the initial rotor configuration.

Our next result is the analogue of Theorem 1.2.1 on regular trees. Call a configuration of rotors *acyclic* if there are no directed cycles of rotors. On a tree, this is equivalent to forbidding directed cycles of length 2: for any pair of neighboring vertices $v \sim w$, if the rotor at v points to w, then the rotor at w does not point to v. As the following result shows, provided we start with an acyclic configuration of rotors, the occupied cluster A_n is a perfect ball for suitable values of n.

Theorem 1.5.1. *Let T be the infinite d-regular tree, and let*

$$B_r = \{x \in T : |x| \leq r\}$$

be the ball of radius r centered at the origin $o \in T$, where $|x|$ is the length of the shortest path from o to x. Write
$$b_r = \#B_r = 1 + d\frac{(d-1)^r - 1}{d - 2}.$$
Let A_n be the region formed by rotor-router aggregation on the infinite d-regular tree, starting from n chips at o. If the initial rotor configuration is acyclic, then

$$A_{b_r} = B_r.$$

The proof of Theorem 1.5.1 uses the *sandpile group* of a finite regular tree with the leaves collapsed to a single vertex. This is an abelian group defined for any graph G whose order is the number of spanning trees of G. In section 4.1 we recall the definition of

the sandpile group and prove the following decomposition theorem expressing the sandpile group of a finite regular tree as a product of cyclic groups.

Theorem 1.5.2. *Let T_n be the regular tree of degree $d = a + 1$ and height n, with leaves collapsed to a single sink vertex and an edge joining the root to the sink. Then writing $\mathbb{Z}_p^q$ for the group $(\mathbb{Z}/p\mathbb{Z}) \oplus \ldots \oplus (\mathbb{Z}/p\mathbb{Z})$ with q summands, the sandpile group of T_n is given by*

$$SP(T_n) \simeq \mathbb{Z}_{1+a}^{a^{n-3}(a-1)} \oplus \mathbb{Z}_{1+a+a^2}^{a^{n-4}(a-1)} \oplus \ldots \oplus \mathbb{Z}_{1+a+\ldots+a^{n-2}}^{a-1} \oplus \mathbb{Z}_{1+a+\ldots+a^{n-1}}.$$

For example, taking $d = 3$ we obtain that the sandpile group of the regular ternary tree of height n has the decomposition

$$SP(T_n) \simeq (\mathbb{Z}_3)^{2^{n-3}} \oplus (\mathbb{Z}_7)^{2^{n-4}} \oplus \ldots \oplus \mathbb{Z}_{2^{n-1}-1} \oplus \mathbb{Z}_{2^n-1}.$$

Toumpakari [47] studied the sandpile group of the ball B_n inside the infinite d-regular tree. Her setup differs slightly from ours in that there is no edge connecting the root to the sink. She found the rank, exponent, and order of $SP(B_n)$ and conjectured a formula for the ranks of its Sylow p-subgroups. We use Theorem 1.5.2 to give a proof of her conjecture in section 4.1.3.

Chen and Schedler [10] study the sandpile group of thick trees (i.e. trees with multiple edges) without collapsing the leaves to the sink. They obtain quite a different product formula in this setting.

In section 4.2 we define the *rotor-router group* of a graph and show that it is isomorphic to the sandpile group. We then use this isomorphism to prove Theorem 1.5.1.

Much previous work on the rotor-router model has taken the form of comparing the behavior of rotor-router walk with the expected behavior of random walk. For example, Cooper and Spencer [12] show that for any configuration of chips on even lattice sites in $\mathbb{Z}^d$, letting each chip perform rotor-router walk for n steps results in a configuration that differs by only constant error from the expected configuration had the chips performed independent random walks. We continue in this vein by investigating the recurrence and transience of rotor-router walk on trees. A walk which never returns to the origin visits each vertex only finitely many times, so the positions of the rotors after a walk has escaped to infinity are well-defined. We construct two "extremal" rotor configurations on the infinite ternary tree, one for which walks exactly alternate returning to the origin with escaping to infinity, and one for which every walk returns to the origin. The latter behavior is something of

a surprise: to our knowledge it represents the first example of rotor-router walk behaving fundamentally differently from the expected behavior of random walk.

In between these two extreme cases, a variety of intermediate behaviors are possible. We say that a binary word $a_1 \ldots a_n$ is an *escape sequence* for the infinite ternary tree if there exists an initial rotor configuration on the tree so that the k-th chip escapes to infinity if and only if $a_k = 1$. The following result characterizes all possible escape sequences on the ternary tree.

Theorem 1.5.3. *Let* $a = a_1 \ldots a_n$ *be a binary word. For* $j \in \{1, 2, 3\}$ *write* $a^{(j)} = a_j a_{j+3} a_{j+6} \ldots$. *Then* a *is an escape sequence for some rotor configuration on the infinite ternary tree if and only if for each* j *and* $k \geq 2$, *every subword of* $a^{(j)}$ *of length* $2^k - 1$ *contains at most* 2^{k-1} *ones.*

Theorem 1.5.3 is proved in section 4.2.3 by expressing the escape sequence corresponding to a rotor configuration on the full tree in terms of the escape sequences of the configurations on each of the principal subtrees.

Chapter 2

Spherical Asymptotics for Point Sources

This chapter is devoted to the proofs of Theorems 1.2.1, 1.2.2 and 1.2.3. In section 2.1, we derive the basic Green's function estimates that are used in the proofs. In section 2.2 we prove the abelian property and Theorem 1.2.2 for the divisible sandpile. In section 2.3 we adapt the methods used for the divisible sandpile to prove Theorem 1.2.3 for the classical abelian sandpile model. Section 2.4 is devoted to the proof of Theorem 1.2.1 for the rotor-router model.

2.1 Basic Estimate

Write $(X_k)_{k \geq 0}$ for simple random walk in $\mathbb{Z}^d$, and for $d \geq 3$ denote by

$$g_1(x, y) = \mathbb{E}_x \#\{k | X_k = y\} \tag{2.1}$$

the expected number of visits to y by simple random walk started at x. This is the *discrete harmonic Green's function* in $\mathbb{Z}^d$. For fixed x, the function $g_1(x, \cdot)$ is harmonic except at x, where its discrete Laplacian is -1. Our notation g_1 is chosen to distinguish between the discrete Green's function in $\mathbb{Z}^d$ and its continuous counterpart g in $\mathbb{R}^d$. For the definition of g, see section 3.1.2. Estimates relating the discrete and continuous Green's functions are discussed in section 3.1.5.

In dimension $d = 2$, simple random walk is recurrent, so the expectation on the

right side of (2.1) infinite. Here we define the *potential kernel*

$$g_1(x, y) = \lim_{n \to \infty} (g_1^n(x, y) - g_1^n(x, x)) \tag{2.2}$$

where

$$g_1^n(x, y) = \mathbb{E}_x \#\{k \leq n \,|\, X_k = y\}.$$

The limit defining g_1 in (2.2) is finite [28, 45], and $g_1(x, \cdot)$ is harmonic except at x, where its discrete Laplacian is -1. Note that (2.2) is the negative of the usual definition of the potential kernel; we have chosen this sign convention so that g_1 has the same Laplacian in dimension two as in higher dimensions.

Fix a real number $m > 0$ and consider the function on $\mathbb{Z}^d$

$$\widetilde{\gamma}_d(x) = |x|^2 + m g_1(o, x). \tag{2.3}$$

Let r be such that $m = \omega_d r^d$, and let

$$\gamma_d(x) = \widetilde{\gamma}_d(x) - \widetilde{\gamma}_d(\lfloor r \rfloor e_1) \tag{2.4}$$

where e_1 is the first standard basis vector in $\mathbb{Z}^d$. The function γ_d plays a central role in our analysis. To see where it comes from, recall the divisible sandpile *odometer function*

$$u(x) = \text{total mass emitted from } x.$$

Let $D_m \subset \mathbb{Z}^d$ be the domain of fully occupied sites for the divisible sandpile formed from a pile of mass m at the origin. For $x \in D_m$, since each neighbor y of x emits an equal amount of mass to each of its $2d$ neighbors, we have

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

$$= \text{mass received by } x - \text{mass emitted by } x$$

$$= 1 - m\delta_{ox}.$$

Moreover, $u = 0$ on ∂D_m, where for $A \subset \mathbb{Z}^d$ we write

$$\partial A = \{x \in A^c \,|\, x \sim y \text{ for some } y \in A\}$$

for the boundary of A. By construction, the function γ_d obeys the same Laplacian condition: $\Delta \gamma_d = 1 - m\delta_o$; and as we will see shortly, $\gamma_d \approx 0$ on ∂B_r. Since we expect the domain D_m

to be close to the ball B_r, we should expect that $u \approx \gamma_d$. In fact, we will first show that u is close to γ_d, and then use this to conclude that D_m is close to B_r.

We will use the following estimates for the Green's function [20, 48]; see also [28, Theorems 1.5.4 and 1.6.2].

$$g_1(x, y) = \begin{cases} -\frac{2}{\pi} \log |x - y| + \kappa + O(|x - y|^{-2}), & d = 2 \\ a_d |x - y|^{2-d} + O(|x - y|^{-d}), & d \geq 3. \end{cases} \tag{2.5}$$

Here $a_d = \frac{2}{(d-2)\omega_d}$, where ω_d is the volume of the unit ball in $\mathbb{R}^d$, and κ is a constant whose value we will not need to know. Here and throughout this thesis, constants in error terms denoted $O(\cdot)$ depend only on d.

We will need an estimate for the function γ_d near the boundary of the ball B_r. We first consider dimension $d = 2$. From (2.5) we have

$$\widetilde{\gamma}_2(x) = \phi(x) - \kappa m + O(m|x|^{-2}), \tag{2.6}$$

where

$$\phi(x) = |x|^2 - \frac{2m}{\pi} \log |x|.$$

In the Taylor expansion of ϕ about $|x| = r$

$$\phi(x) = \phi(r) - \phi'(r)(r - |x|) + \frac{1}{2}\phi''(t)(r - |x|)^2 \tag{2.7}$$

the linear term vanishes, leaving

$$\gamma_2(x) = \left(1 + \frac{m}{\pi t^2}\right)(r - |x|)^2 + O(m|x|^{-2}) \tag{2.8}$$

for some t between $|x|$ and r.

In dimensions $d \geq 3$, from (2.5) we have

$$\widetilde{\gamma}_d(x) = |x|^2 + a_d m |x|^{2-d} + O(m|x|^{-d}).$$

Setting $\phi(x) = |x|^2 + a_d m |x|^{2-d}$, the linear term in the Taylor expansion (2.7) of ϕ about $|x| = r$ again vanishes, yielding

$$\gamma_d(x) = \left(1 + (d-1)(r/t)^d\right)(r - |x|)^2 + O(m|x|^{-d})$$

for t between $|x|$ and r. Together with (2.8), this yields the following estimates in all dimensions $d \geq 2$.

Lemma 2.1.1. *Let γ_d be given by (2.4). For all $x \in \mathbb{Z}^d$ we have*

$$\gamma_d(x) \geq (r - |x|)^2 + O\left(\frac{r^d}{|x|^d}\right). \tag{2.9}$$

Lemma 2.1.2. *Let γ_d be given by (2.4). Then uniformly in r,*

$$\gamma_d(x) = O(1), \qquad x \in B_{r+1} - B_{r-1}.$$

The following lemma is useful for x near the origin, where the error term in (2.9) blows up.

Lemma 2.1.3. *Let γ_d be given by (2.4). Then for sufficiently large r, we have*

$$\gamma_d(x) \geq \frac{r^2}{4}, \qquad \forall x \in B_{r/3}.$$

Proof. Since $\gamma_d(x) - |x|^2$ is superharmonic, it attains its minimum in $B_{r/3}$ at a point z on the boundary. Thus for any $x \in B_{r/3}$

$$\gamma_d(x) - |x|^2 \geq \gamma_d(z) - |z|^2,$$

hence by Lemma 2.1.1

$$\gamma_d(x) \geq (2r/3)^2 - (r/3)^2 + O(1) > \frac{r^2}{4}. \quad \square$$

Lemmas 2.1.1 and 2.1.3 together imply the following.

Lemma 2.1.4. *Let γ_d be given by (2.4). There is a constant a depending only on d, such that $\gamma_d \geq -a$ everywhere.*

2.2 Divisible Sandpile

2.2.1 Abelian Property

In this section we prove the abelian property of the divisible sandpile mentioned in the introduction. We work in continuous time. Fix $\tau > 0$, and let $T : [0, \tau] \to \mathbb{Z}^d$ be a function having only finitely many discontinuities in the interval $[0, t]$ for every $t < \tau$. The

value $T(t)$ represents the site being toppled at time t. The odometer function at time t is given by

$$u_t(x) = \mathcal{L}\left(T^{-1}(x) \cap [0, t]\right),$$

where $\mathcal{L}$ denotes Lebesgue measure. We will say that T is a *legal toppling function* for an initial configuration ν_0 if for every $0 \le t \le \tau$

$$\nu_t(T(t)) \ge 1$$

where

$$\nu_t(x) = \nu_0(x) + \Delta u_t(x) \tag{2.10}$$

is the amount of mass present at x at time t. If in addition $\nu_\tau \le 1$, we say that T is *complete*. The abelian property can now be stated as follows.

Lemma 2.2.1. *If $T_1 : [0, \tau_1] \to \mathbb{Z}^d$ and $T_2 : [0, \tau_2] \to \mathbb{Z}^d$ are complete legal toppling functions for an initial configuration ν_0, then for any $x \in \mathbb{Z}^d$*

$$\mathcal{L}\left(T_1^{-1}(x)\right) = \mathcal{L}\left(T_2^{-1}(x)\right).$$

In particular, $\tau_1 = \tau_2$ and the final configurations ν_{τ_1}, ν_{τ_2} are identical.

Proof. For $i = 1, 2$ write

$$u_t^i(x) = \mathcal{L}\left(T_i^{-1}(x) \cap [0, t]\right);$$
$$\nu_t^i(x) = \nu_0(x) + \Delta u_t^i(x).$$

Write $u^i = u_{\tau_i}^i$ and $\nu^i = \nu_{\tau_i}^i$. Let $t(0) = 0$ and let $t(1) < t(2) < \dots$ be the points of discontinuity for T_1. Let x_k be the value of T_1 on the interval $(t(k-1), t(k))$. We will show by induction on k that

$$u^2(x_k) \ge u_{t(k)}^1(x_k). \tag{2.11}$$

Note that for any $x \ne x_k$, if $u_{t(k)}^1(x) > 0$, then letting $j < k$ be maximal such that $x_j = x$, since $T_1 \ne x$ on the interval $(t(j), t(k))$, it follows from the inductive hypothesis (2.11) that

$$u^2(x) = u^2(x_j) \ge u_{t(j)}^1(x_j) = u_{t(k)}^1(x). \tag{2.12}$$

Since T_1 is legal and T_2 is complete, we have

$$\nu^2(x_k) \le 1 \le \nu_{t(k)}^1(x_k)$$

hence

$$\Delta u^2(x_k) \leq \Delta u^1_{t(k)}(x_k).$$

Since T_1 is constant on the interval $(t(k-1), t(k))$ we obtain

$$u^2(x_k) \geq u^1_{t(k)}(x_k) + \frac{1}{2d} \sum_{x \sim x_k} (u^2(x) - u^1_{t(k-1)}(x)).$$

By (2.12), each term in the sum on the right side is nonnegative, completing the inductive step.

Since $t(k) \uparrow \tau_1$ as $k \uparrow \infty$, the right side of (2.12) converges to $u^1(x)$ as $k \uparrow \infty$, hence $u^2 \geq u^1$. After interchanging the roles of T_1 and T_2, the result follows. $\qquad\square$

2.2.2 Proof of Theorem 1.2.2

Recall that a function s on $\mathbb{Z}^d$ is *superharmonic* if $\Delta s \leq 0$. Given a function γ on $\mathbb{Z}^d$ the *least superharmonic majorant* of γ is the function

$$s(x) = \inf\{f(x) \mid f \text{ is superharmonic and } f \geq \gamma\}.$$

Note that if f is superharmonic and $f \geq \gamma$ then

$$f(x) \geq \frac{1}{2d} \sum_{y \sim x} f(y) \geq \frac{1}{2d} \sum_{y \sim x} s(y).$$

Taking the infimum on the left side we obtain that s is superharmonic.

Lemma 2.2.2. *Let $T : [0, \tau] \to \mathbb{Z}^d$ be a complete legal toppling function for the initial configuration ν_0, and let*

$$u(x) = \mathcal{L}(T^{-1}(x))$$

be the corresponding odometer function for the divisible sandpile. Then $u = s + \gamma$, where

$$\gamma(x) = |x|^2 + \sum_{y \in \mathbb{Z}^d} g_1(x, y)\nu_0(y)$$

and s is the least superharmonic majorant of $-\gamma$.

Proof. From (2.10) we have

$$\Delta u = \nu_\tau - \nu_0 \leq 1 - \nu_0.$$

Since $\Delta\gamma = 1 - \nu_0$, the difference $u - \gamma$ is superharmonic. As u is nonnegative, it follows that $u - \gamma \geq s$. For the reverse inequality, note that $s + \gamma - u$ is superharmonic on the domain $D = \{x \mid \nu_\tau(x) = 1\}$ of fully occupied sites and is nonnegative outside D, hence nonnegative inside D as well. $\qquad\square$

We now turn to the case of a point source mass m started at the origin: $\nu_0 = m\delta_o$. More general starting densities are treated in Chapter 3. In the case of a point source of mass m, the natural question is to identify the shape of the resulting domain D_m of fully occupied sites, i.e. sites x for which $\mu(x) = 1$. According to Theorem 2.2.3, D_m is extremely close to a ball of volume m; in fact, the error in the radius is a constant independent of m. As before, for $r \geq 0$ we write

$$B_r = \{x \in \mathbb{Z}^d \ : \ |x| < r\}$$

for the lattice ball of radius r centered at the origin.

Theorem 2.2.3. *For $m \geq 0$ let $D_m \subset \mathbb{Z}^d$ be the domain of fully occupied sites for the divisible sandpile formed from a pile of size m at the origin. There exist constants c, c' depending only on d, such that*

$$B_{r-c} \subset D_m \subset B_{r+c'},$$

where $r = (m/\omega_d)^{1/d}$ and ω_d is the volume of the unit ball in $\mathbb{R}^d$.

The idea of the proof is to use Lemma 2.2.2 along with the basic estimates on γ, Lemmas 2.1.1 and 2.1.2, to obtain estimates on the odometer function

$$u(x) = \text{total mass emitted from } x.$$

We will need the following simple observation.

Lemma 2.2.4. *For every point $x \in D_m - \{o\}$ there is a path $x = x_0 \sim x_1 \sim \ldots \sim x_k = o$ in D_m with $u(x_{i+1}) \geq u(x_i) + 1$.*

Proof. If $x_i \in D_m - \{o\}$, let x_{i+1} be a neighbor of x_i maximizing $u(x_{i+1})$. Then $x_{i+1} \in D_m$ and

$$u(x_{i+1}) \geq \frac{1}{2d} \sum_{y \sim x_i} u(y)$$
$$= u(x_i) + \Delta u(x_i)$$
$$= u(x_i) + 1,$$

where in the last step we have used the fact that $x_i \in D_m$. $\qquad\square$

Proof of Theorem 2.2.3. We first treat the inner estimate. Let γ_d be given by (2.4). By Lemma 2.2.2 the function $u - \gamma_d$ is superharmonic, so its minimum in the ball B_r is attained on the boundary. Since $u \geq 0$, we have by Lemma 2.1.2

$$u(x) - \gamma_d(x) \geq O(1), \qquad x \in \partial B_r.$$

Hence by Lemma 2.1.1,

$$u(x) \geq (r - |x|)^2 + O(r^d / |x|^d), \qquad x \in B_r. \tag{2.13}$$

It follows that there is a constant c, depending only on d, such that $u(x) > 0$ whenever $r/3 \leq |x| < r - c$. Thus $B_{r-c} - B_{r/3} \subset D_m$. For $x \in B_{r/3}$, by Lemma 2.1.3 we have $u(x) > r^2/4 + O(1) > 0$, hence $B_{r/3} \subset D_m$.

For the outer estimate, note that $u - \gamma_d$ is harmonic on D_m. By Lemma 2.1.4 we have $\gamma_d \geq -a$ everywhere, where a depends only on d. Since u vanishes on ∂D_m it follows that $u - \gamma_d \leq a$ on D_m. Now for any $x \in D_m$ with $r - 1 < |x| \leq r$, we have by Lemma 2.1.2

$$u(x) \leq \gamma_d(x) + a \leq c'$$

for a constant c' depending only on d. Lemma 2.2.4 now implies that $D_m \subset B_{r+c'+1}$. $\quad\square$

2.3   Classical Sandpile

We consider a generalization of the classical abelian sandpile, proposed by Fey and Redig [18]. Each site in $\mathbb{Z}^d$ begins with a "hole" of depth H. Thus, each site absorbs the first H grains it receives, and thereafter functions normally, toppling once for each additional $2d$ grains it receives. If H is negative, we can interpret this as saying that every site starts with $h = -H$ grains of sand already present. Aggregation is only well-defined in the regime $h \leq 2d - 2$, since for $h = 2d - 1$ the addition of a single grain already causes every site in $\mathbb{Z}^d$ to topple infinitely often.

Let $S_{n,H}$ be the set of sites that are visited if n particles start at the origin in $\mathbb{Z}^d$. Fey and Redig [18, Theorem 4.7] prove that

$$\lim_{H \to \infty} \limsup_{n \to \infty} \frac{H}{n} \# \left(S_{n,H} \triangle B_{H^{-1/d} r} \right) = 0,$$

where $n = \omega_d r^d$, and $\triangle$ denotes symmetric difference. The following theorem strengthens this result.

Theorem 2.3.1. *Fix an integer $H \geq 2 - 2d$. Let $S_n = S_{n,H}$ be the set of sites that are visited by the classical abelian sandpile model in $\mathbb{Z}^d$, starting from n particles at the origin, if every lattice site begins with a hole of depth H. Write $n = \omega_d r^d$. Then*

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

where

$$c_1 = (2d - 1 + H)^{-1/d}$$

and c_2 is a constant depending only on d. Moreover if $H \geq 1 - d$, then for any $\epsilon > 0$ we have

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

where

$$c_1' = (d - \epsilon + H)^{-1/d}$$

and c_2' is independent of n but may depend on d, H and ϵ.

Note that the ratio $c_1/c_1' \uparrow 1$ as $H \uparrow \infty$. Thus, the classical abelian sandpile run from an initial state in which each lattice site starts with a deep hole yields a shape very close to a ball. Intuitively, one can think of the classical sandpile with deep holes as approximating the divisible sandpile, whose limiting shape is a ball by Theorem 2.2.3. Following this intuition, we can adapt the proof of Theorem 2.2.3 to prove Theorem 2.3.1; just one additional averaging trick is needed, which we explain below.

Consider the odometer function for the abelian sandpile

$$u(x) = \text{total number of grains emitted from } x.$$

Let $T_n = \{x | u(x) > 0\}$ be the set of sites which topple at least once. Then

$$T_n \subset S_n \subset T_n \cup \partial T_n.$$

In the final state, each site which has toppled retains between 0 and $2d-1$ grains, in addition to the H that it absorbed. Hence

$$H \leq \Delta u(x) + n\delta_{ox} \leq 2d - 1 + H, \qquad x \in T_n. \tag{2.14}$$

We can improve the lower bound by averaging over a small box. For $x \in \mathbb{Z}^d$ let

$$Q(x, k) = \{y \in \mathbb{Z}^d : ||x - y||_\infty \leq k\}$$

be the box of side length $2k + 1$ centered at x, and let

$$u^{(k)}(x) = (2k + 1)^{-d} \sum_{y \in Q(x,k)} u(y).$$

Write

$$T_n^{(k)} = \{x \mid Q(x, k) \subset T_n\}.$$

Le Borgne and Rossin [31] observe that if T is a set of sites all of which topple, the number of grains remaining in T is at least the number of edges internal to T: indeed, for each internal edge, the endpoint that topples last sends the other a grain which never moves again. Since the box $Q(x, k)$ has $2dk(2k + 1)^{d-1}$ internal edges, we have

$$\Delta u^{(k)}(x) \geq (d + H) \frac{2k}{2k + 1}, \qquad x \in T_n^{(k)}. \tag{2.15}$$

The following lemma is analogous to Lemma 2.2.4.

Lemma 2.3.2. *For every point $x \in T_n$ adjacent to ∂T_n there is a path $x = x_0 \sim x_1 \sim \ldots \sim x_m = o$ in T_n with $u(x_{i+1}) \geq u(x_i) + 1$.*

Proof. By (2.14) we have

$$\frac{1}{2d} \sum_{y \sim x_i} u(y) \geq u(x_i).$$

Since $u(x_{i-1}) < u(x_i)$, some term $u(y)$ in the sum above must exceed $u(x_i)$. Let $x_{i+1} = y$. $\qquad \square$

Proof of Theorem 2.3.1. Let

$$\widetilde{\xi}_d(x) = (2d - 1 + H)|x|^2 + ng_1(o, x),$$

and let

$$\xi_d(x) = \widetilde{\xi}_d(x) - \widetilde{\xi}_d(\lfloor c_1 r \rfloor e_1).$$

Taking $m = n/(2d - 1 + H)$ in Lemma 2.1.2, we have

$$u(x) - \xi_d(x) \geq -\xi_d(x) = O(1), \qquad x \in \partial B_{c_1 r}. \tag{2.16}$$

By (2.14), $u - \xi_d$ is superharmonic, so (2.16) holds in all of $B_{c_1 r}$. Hence by Lemma 2.1.1

$$u(x) \geq \frac{(r - |x|)^2}{2d - 1 + H} + O(r^d / |x|^d), \qquad x \in B_{c_1 r}. \tag{2.17}$$

It follows that u is positive on $B_{c_1 r - c_2} - B_{c_1 r/3}$ for a suitable constant c_2. For $x \in B_{c_1 r/3}$, by Lemma 2.1.3 we have $u(x) > c_1^2 r^2/4 + O(1) > 0$. Thus $B_{c_1 r - c_2} \subset T_n \subset S_n$.

For the outer estimate, let

$$\hat{\psi}_d(x) = (d - \epsilon + H)|x|^2 + n g_1(o, x).$$

Choose k large enough so that $(d + H)\frac{2k}{2k+1} \geq d - \epsilon + H$ and define

$$\widetilde{\psi}_d(x) = (2k + 1)^{-d} \sum_{y \in Q(x,k)} \hat{\psi}_d(y).$$

Finally, let

$$\psi_d(x) = \widetilde{\psi}_d(x) - \widetilde{\psi}_d(\lfloor c_1' r \rfloor e_1).$$

By (2.15), $u^{(k)} - \psi_d$ is subharmonic on $T_n^{(k)}$. Taking $m = n/(d - \epsilon + H)$ in Lemma 2.1.4, there is a constant a such that $\psi_d \geq -a$ everywhere. Since $u^{(k)} \leq (2d + H)^{(d+1)k}$ on $\partial T_n^{(k)}$ it follows that $u^{(k)} - \psi_d \leq a + (2d + H)^{(d+1)k}$ on $T_n^{(k)}$. Now for any $x \in S_n$ with $c_1' r - 1 < |x| \leq c_1' r$ we have by Lemma 2.1.2

$$u^{(k)}(x) \leq \psi_d(x) + a + (2d + H)^{(d+1)k} \leq \tilde{c}_2$$

for a constant $\tilde{c}_2$ depending only on d, H and ϵ. Then $u(x) \leq c_2' := (2k+1)^d \tilde{c}_2$. Lemma 2.3.2 now implies that $T_n \subset B_{c_1' r + c_2'}$, and hence

$$S_n \subset T_n \cup \partial T_n \subset B_{c_1' r + c_2' + 1}. \quad \square$$

We remark that the crude bound of $(2d + H)^{(d+1)k}$ used in the proof of the outer estimate can be improved to a bound of order $k^2 H$, and the final factor of $(2k + 1)^d$ can be replaced by a constant factor independent of k and H, using the fact that a nonnegative function on $\mathbb{Z}^d$ with bounded Laplacian cannot grow faster than quadratically; see Lemma 3.1.18.

2.4 Rotor-Router Model

Given a function f on $\mathbb{Z}^d$, for a directed edge (x, y) write

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

Given a function s on directed edges in $\mathbb{Z}^d$, write

$$\operatorname{div} s(x) = \frac{1}{2d} \sum_{y \sim x} s(x, y).$$

The discrete Laplacian of f is then given by

$$\Delta f(x) = \operatorname{div} \nabla f = \frac{1}{2d} \sum_{y \sim x} f(y) - f(x).$$

2.4.1   Inner Estimate

Fixing $n \geq 1$, consider the odometer function for rotor-router aggregation

$$u(x) = \text{total number exits from } x \text{ by the first } n \text{ particles.}$$

We learned the idea of using the odometer function to study the rotor-router shape from Matt Cook [11].

Lemma 2.4.1. *For a directed edge (x, y) in $\mathbb{Z}^d$, denote by $\kappa(x, y)$ the net number of crossings from x to y performed by the first n particles in rotor-router aggregation. Then*

$$\nabla u(x, y) = -2d\kappa(x, y) + R(x, y) \tag{2.18}$$

for some edge function R which satisfies

$$|R(x, y)| \leq 4d - 2$$

for all edges (x, y).

Remark. In the more general setting of rotor stacks of bounded discrepancy, the $4d - 2$ will be replaced by a different constant here.

Proof. Writing $N(x, y)$ for the number of particles routed from x to y, we have

$$\frac{u(x) - 2d + 1}{2d} \leq N(x, y) \leq \frac{u(x) + 2d - 1}{2d}$$

hence

$$|\nabla u(x, y) + 2d\kappa(x, y)| = |u(y) - u(x) + 2dN(x, y) - 2dN(y, x)|$$
$$\leq 4d - 2. \quad \square$$

For the remainder of this section $C_0, C_1, \ldots$ denote constants depending only on d.

Lemma 2.4.2. *Let $\Omega \subset \mathbb{Z}^d - \{o\}$ be a finite set. Then*

$$\sum_{y \in \Omega} |y|^{1-d} \leq C_0 \operatorname{Diam}(\Omega).$$

Proof. Let

$$\mathcal{S}_k = \{y \in \mathbb{Z}^d \ : \ k \leq |y| < k+1\}.$$

Then

$$\sum_{y \in S_k} |y|^{1-d} \leq k^{1-d} \# \mathcal{S}_k \leq C_0.$$

Since Ω can intersect at most $\operatorname{Diam}(\Omega) + 1$ distinct sets $\mathcal{S}_k$, the proof is complete. $\qquad \square$

Lemma 2.4.3. *Let $G = G_{B_r}$ be the Green's function for simple random walk in $\mathbb{Z}^d$ stopped on exiting B_r. Let $x \in B_r$ with $|x| > r - \rho$. Then*

$$\sum_{\substack{y \in B_r \\ |x-y| < 3\rho}} \sum_{z \sim y} |G(x, y) - G(x, z)| \leq C_1 \rho. \tag{2.19}$$

Proof. Let $(X_t)_{t \geq 0}$ denote simple random walk in $\mathbb{Z}^d$, and let T be the first exit time from B_r. For fixed y, the function

$$A(x) = g_1(x, y) - \mathbb{E}_x g_1(X_T, y) \tag{2.20}$$

has Laplacian $\Delta A(x) = -\delta_{xy}$ in B_r and vanishes on ∂B_r, hence $A(x) = G(x, y)$. Let $x, y \in B_r$ and $z \sim y$. From (2.5) we have

$$|g_1(x, y) - g_1(x, z)| \leq \frac{C_2}{|x-y|^{d-1}}, \qquad y, z \neq x.$$

Using the triangle inequality together with (2.20), we obtain

$$|G(x, y) - G(x, z)| \leq |g_1(x, y) - g_1(x, z)| + \mathbb{E}_x |g_1(X_T, y) - g_1(X_T, z)|$$

$$\leq \frac{C_2}{|x-y|^{d-1}} + \sum_{w \in \partial B_r} H_x(w) \frac{C_2}{|w-y|^{d-1}},$$

where $H_x(w) = \mathbb{P}_x(X_T = w)$.

Write $D = \{y \in B_r \ : \ |x-y| < 3\rho\}$. Then

$$\sum_{\substack{y \in D \\ y \neq x}} \sum_{\substack{z \sim y \\ z \neq x}} |G(x, y) - G(x, z)| \leq C_3 \rho + C_2 \sum_{w \in \partial B_r} H_x(w) \sum_{y \in D} |w-y|^{1-d}. \tag{2.21}$$
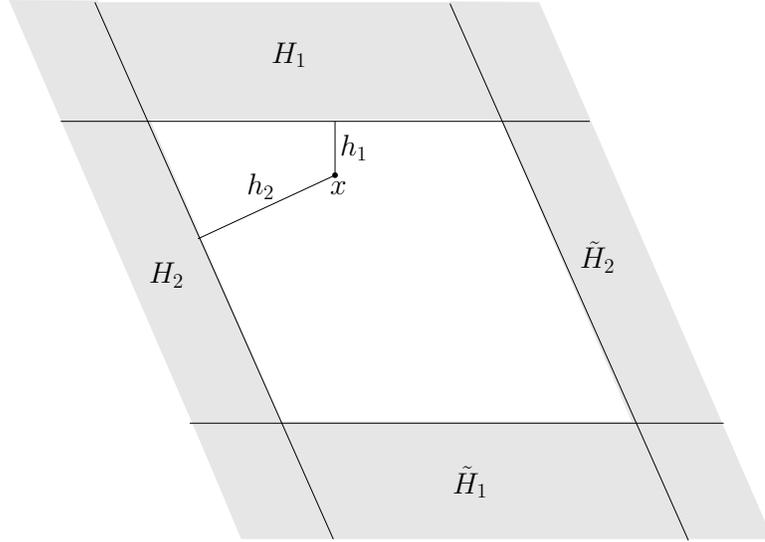
Figure 2.1: Diagram for the Proof of Lemma 2.4.4.

By Lemma 2.4.2, the inner sum on the right is at most $C_0 \operatorname{Diam}(D) = 6C_0\rho$, so the right side of (2.21) is bounded above by $C_1\rho$ for a suitable C_1.

Finally, the terms in which y or z coincides with x make a negligible contribution to the sum in (2.19), since for $y \sim x \in \mathbb{Z}^d$

$$|G(x,x) - G(x,y)| \leq |g_1(x,x) - g_1(x,y)| + \mathbb{E}_x|g_1(X_T,x) - g_1(X_T,y)| \leq C_4. \quad \square$$

Lemma 2.4.4. *Let H_1, H_2 be linear half-spaces in $\mathbb{Z}^d$, not necessarily parallel to the coordinate axes. Let T_i be the first hitting time of H_i. If $x \notin H_1 \cup H_2$, then*

$$\mathbb{P}_x(T_1 > T_2) \leq \frac{5}{2}\frac{h_1 + 1}{h_2}\left(1 + \frac{1}{2h_2}\right)^2$$

where h_i is the distance from x to H_i.

Proof. If one of H_1, H_2 contains the other, the result is vacuous. Otherwise, let $\widetilde{H}_i$ be the half-space shifted parallel to H_i^c by distance $2h_2$ in the direction of x, and let $\widetilde{T}_i$ be the first hitting time of $H_i \cup \widetilde{H}_i$. Let $(X_t)_{t\geq 0}$ denote simple random walk in $\mathbb{Z}^d$, and write M_t for the (signed) distance from X_t to the hyperplane defining the boundary of H_1, with

$M_0 = h_1$. Then M_t is a martingale with bounded increments. Since $\mathbb{E}_x \widetilde{T}_1 < \infty$, we obtain from optional stopping

$$h_1 = \mathbb{E}_x M_{\widetilde{T}_1} \geq 2h_2 \, \mathbb{P}_x \left(X_{\widetilde{T}_1} \in \widetilde{H}_1 \right) - \mathbb{P}_x(X_{\widetilde{T}_1} \in H_1),$$

hence

$$\mathbb{P}_x \left(X_{\widetilde{T}_1} \in \widetilde{H}_1 \right) \leq \frac{h_1 + 1}{2h_2}. \tag{2.22}$$

Likewise, $dM_t^2 - t$ is a martingale with bounded increments, giving

$$\mathbb{E}_x \widetilde{T}_1 \leq d \, \mathbb{E}_x \, M_{\widetilde{T}_1}^2$$

$$\leq d(2h_2 + 1)^2 \, \mathbb{P}_x \left(X_{\widetilde{T}_1} \in \widetilde{H}_1 \right)$$

$$\leq d(h_1 + 1)(2h_2 + 1) \left(1 + \frac{1}{2h_2} \right). \tag{2.23}$$

Let $T = \min(\widetilde{T}_1, \widetilde{T}_2)$. Denoting by D_t the distance from X_t to the hyperplane defining the boundary of H_2, the quantity

$$N_t = \frac{d}{2} \left(D_t^2 + (2h_2 - D_t)^2 \right) - t$$

is a martingale. Writing $p = \mathbb{P}_x(T = \widetilde{T}_2)$ we have

$$dh_2^2 = \mathbb{E}N_0 = \mathbb{E}N_T \geq p\frac{d}{2}(2h_2)^2 + (1-p)dh_2^2 - \mathbb{E}_x T$$

$$\geq (1 + p)dh_2^2 - \mathbb{E}_x T$$

hence by (2.23)

$$p \leq \frac{\mathbb{E}_x T}{dh_2^2} \leq 2\frac{h_1 + 1}{h_2} \left(1 + \frac{1}{2h_2} \right)^2.$$

Finally by (2.22)

$$\mathbb{P}(T_1 > T_2) \leq p + \mathbb{P}(X_{\widetilde{T}_1} \in \widetilde{H}_1) \leq \frac{5}{2} \frac{h_1 + 1}{h_2} \left(1 + \frac{1}{2h_2} \right)^2. \quad \square$$

Lemma 2.4.5. Let $x \in B_r$ and let $\rho = r + 1 - |x|$. Let

$$\mathcal{S}_k^* = \{y \in B_r \ : \ 2^k \rho < |x - y| \leq 2^{k+1}\rho\}. \tag{2.24}$$

Let τ_k be the first hitting time of $\mathcal{S}_k^*$, and T the first exit time from B_r. Then

$$\mathbb{P}_x(\tau_k < T) \leq C_2 2^{-k}.$$
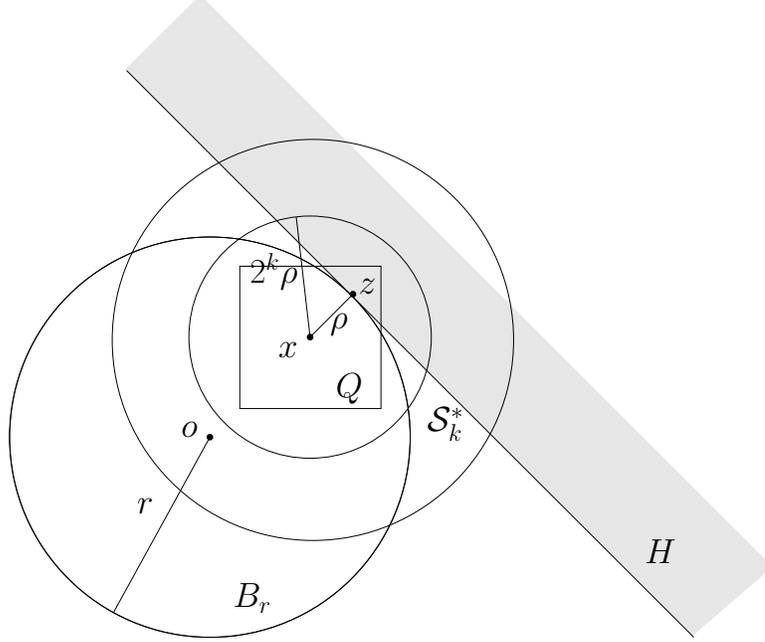
Figure 2.2: Diagram for the proof of Lemma 2.4.5.

Proof. Let H be the outer half-space tangent to B_r at the point $z \in \partial B_r$ closest to x. Let Q be the cube of side length $2^k \rho / \sqrt{d}$ centered at x. Then Q is disjoint from $\mathcal{S}_k^*$, hence

$$\mathbb{P}_x(\tau_k < T) \leq \mathbb{P}_x(T_{\partial Q} < T) \leq \mathbb{P}_x(T_{\partial Q} < T_H)$$

where $T_{\partial Q}$ and T_H are the first hitting times of ∂Q and H. Let $H_1, \ldots, H_{2d}$ be the outer half-spaces defining the faces of Q, so that $Q = H_1^c \cap \ldots \cap H_{2d}^c$. By Lemma 2.4.4 we have

$$\mathbb{P}_x(T_{\partial Q} < T_H) \leq \sum_{i=1}^{2d} \mathbb{P}_x(T_{H_i} < T_H)$$

$$\leq \frac{5}{2} \sum_{i=1}^{2d} \frac{\operatorname{dist}(x, H) + 1}{\operatorname{dist}(x, H_i)} \left(1 + \frac{1}{2\operatorname{dist}(x, H_i)}\right)^2.$$

Since $\operatorname{dist}(x, H) = |x - z| \leq \rho$ and $\operatorname{dist}(x, H_i) = 2^{k-1}\rho/\sqrt{d}$, and $\rho \geq 1$, taking $C_2 = 20\, d^{3/2}(1 + \sqrt{d})^2$ completes the proof. $\qquad \square$

Lemma 2.4.6. *Let $G = G_{B_r}$ be the Green's function for random walk stopped on exiting B_r. Let $x \in B_r$ and let $\rho = r - |x|$. Then*

$$\sum_{y \in B_r} \sum_{z \sim y} |G(x, y) - G(x, z)| \leq C_3 \rho \log \frac{r}{\rho}.$$

Proof. Let $\mathcal{S}_k^*$ be given by (2.24), and let

$$W = \{w \in \partial(\mathcal{S}_k^* \cup \partial \mathcal{S}_k^*) \ : \ |w - x| < 2^k \rho\}.$$

Let τ_W be the first hitting time of W and T the first exit time from B_r. For $w \in W$ let

$$H_x(w) = \mathbb{P}_x(X_{\tau_W \wedge T} = w).$$

Fixing $y \in \mathcal{S}_k^*$ and $z \sim y$, simple random walk started at x must hit W before hitting either y or z, hence

$$|G(x,y) - G(x,z)| \leq \sum_{w \in W} H_x(w)|G(w,y) - G(w,z)|.$$

For any $w \in W$ we have $|w| > r - (2^k + 1)\rho$ and $|y - w| < 3 \cdot 2^k \rho$. Lemma 2.4.3 yields

$$\sum_{y \in \mathcal{S}_k^*} \sum_{z \sim y} |G(x,y) - G(x,z)| \leq C_1(2^k + 1)\rho \sum_{w \in W} H_x(w).$$

By Lemma 2.4.5 we have $\sum_{w \in W} H_x(w) \leq C_2 2^{-k}$, so the above sum is at most $2C_1 C_2 \rho$. Since the union of shells $\mathcal{S}_0^*, \mathcal{S}_1^*, \ldots, \mathcal{S}_{\lceil \log_2(r/\rho) \rceil}^*$ covers all of B_r except for those points within distance ρ of x, and $\sum_{|x-y| \leq \rho} \sum_{z \sim y} |G(x,y) - G(x,z)| \leq C_1 \rho$ by Lemma 2.4.3, the result follows. $\qquad\square$

Proof of Theorem 1.2.1, Inner Estimate. Let κ and R be defined as in Lemma 2.4.1. Since the net number of particles to enter a site $x \neq o$ is at most one, we have $2d \operatorname{div} \kappa(x) \geq -1$. Likewise $2d \operatorname{div} \kappa(o) = n - 1$. Taking the divergence in (2.18), we obtain

$$\Delta u(x) \leq 1 + \operatorname{div} R(x), \qquad x \neq o; \tag{2.25}$$

$$\Delta u(o) = 1 - n + \operatorname{div} R(o). \tag{2.26}$$

Let T be the first exit time from B_r, and define

$$f(x) = \mathbb{E}_x u(X_T) - \mathbb{E}_x T + n \, \mathbb{E}_x \, \#\{j < T | X_j = 0\}.$$

Then $\Delta f(x) = 1$ for $x \in B_r - \{o\}$ and $\Delta f(o) = 1 - n$. Moreover $f \geq 0$ on ∂B_r. It follows from Lemma 2.1.2 with $m = n$ that $f \geq \gamma - C_4$ on B_r for a suitable constant C_4.

We have

$$u(x) - \mathbb{E}_x u(X_T) = \sum_{k \geq 0} \mathbb{E}_x \Big(u(X_{k \wedge T}) - u(X_{(k+1) \wedge T}) \Big).$$

Each summand on the right side is zero on the event $\{T \le k\}$, hence

$$\mathbb{E}_x\Big(u(X_{k\wedge T}) - u(X_{(k+1)\wedge T}) \mid \mathcal{F}_{k\wedge T}\Big) = -\Delta u(X_k)1_{\{T>k\}}.$$

Taking expectations and using (2.25) and (2.26), we obtain

$$u(x) - \mathbb{E}_x u(X_T) \ge \sum_{k\ge 0} \mathbb{E}_x\Big[1_{\{T>k\}}(n1_{\{X_k=o\}} - 1 - \operatorname{div} R(X_k))\Big]$$

$$= n\,\mathbb{E}_x\,\#\{k < T | X_k = o\} - \mathbb{E}_x T - \sum_{k\ge 0}\mathbb{E}_x\Big[1_{\{T>k\}}\operatorname{div} R(X_k)\Big],$$

hence

$$u(x) - f(x) \ge -\frac{1}{2d}\sum_{k\ge 0}\mathbb{E}_x\left[1_{\{T>k\}}\sum_{z\sim X_k} R(X_k,z)\right]. \tag{2.27}$$

Since random walk exits B_r with probability at least $\frac{1}{2d}$ every time it reaches a site adjacent to the boundary ∂B_r, the expected time spent adjacent to the boundary before time T is at most $2d$. Since $|R| \le 4d$, the terms in (2.27) with $z \in \partial B_r$ contribute at most $16d^3$ to the sum. Thus

$$u(x) - f(x) \ge -\frac{1}{2d}\sum_{k\ge 0}\mathbb{E}_x\left[\sum_{\substack{y,z\in B_r \\ y\sim z}} 1_{\{T>k\}\cap\{X_k=y\}}R(y,z)\right] - 8d^2.$$

For $y \in B_r$ we have $\{X_k = y\} \cap \{T > k\} = \{X_{k\wedge T} = y\}$, hence

$$u(x) - f(x) \ge -\frac{1}{2d}\sum_{k\ge 0}\sum_{\substack{y,z\in B_r \\ y\sim z}} \mathbb{P}_x(X_{k\wedge T} = y)R(y,z) - 8d^2. \tag{2.28}$$

Write $p_k(y) = \mathbb{P}_x(X_{k\wedge T} = y)$. Note that since ∇f and κ are antisymmetric, R is antisymmetric. Thus

$$\sum_{\substack{y,z\in B_r \\ y\sim z}} p_k(y)R(y,z) = -\sum_{\substack{y,z\in B_r \\ y\sim z}} p_k(z)R(y,z)$$

$$= \sum_{\substack{y,z\in B_r \\ y\sim z}} \frac{p_k(y) - p_k(z)}{2}R(y,z).$$

Summing over k and using the fact that $|R| \le 4d$, we conclude from (2.28) that

$$u(x) \ge f(x) - \sum_{\substack{y,z\in B_r \\ y\sim z}} |G(x,y) - G(x,z)| - 8d^2,$$

where $G = G_{B_r}$ is the Green's function for simple random walk stopped on exiting B_r. By Lemma 2.4.6 we obtain

$$u(x) \geq f(x) - C_3(r - |x|) \log \frac{r}{r - |x|} - 8d^2.$$

Using the fact that $f \geq \gamma - C_4$, we obtain from Lemma 2.1.1

$$u(x) \geq (r - |x|)^2 - C_3(r - |x|) \log \frac{r}{r - |x|} + O\left(\frac{r^d}{|x|^d}\right).$$

The right side is positive provided $r/3 \leq |x| < r - C_5 \log r$. For $x \in B_{r/3}$, by Lemma 2.1.3 we have $u(x) > r^2/4 - C_3 r \log \frac{3}{2} > 0$, hence $B_{r - C_5 \log r} \subset A_n$. □

2.4.2 Outer Estimate

The following result is due to Holroyd and Propp (unpublished); we include a proof for the sake of completeness. Notice that the bound in (2.29) does not depend on the number of particles.

Proposition 2.4.7. *Let Γ be a finite connected graph, and let $Y \subset Z$ be subsets of the vertex set of Γ. Let s be a nonnegative integer-valued function on the vertices of Γ. Let $H_w(s, Y)$ be the expected number of particles stopping in Y if $s(x)$ particles start at each vertex x and perform independent simple random walks stopped on first hitting Z. Let $H_r(s, Y)$ be the number of particles stopping in Y if $s(x)$ particles start at each vertex x and perform rotor-router walks stopped on first hitting Z. Let $H(x) = H_w(1_x, Y)$. Then*

$$|H_r(s, Y) - H_w(s, Y)| \leq \sum_{u \in G} \sum_{v \sim u} |H(u) - H(v)| \tag{2.29}$$

independent of s and the initial positions of the rotors.

Proof. For each vertex u, arbitrarily choose a neighbor $\eta(u)$. Order the neighbors $\eta(u) = v_1, v_2, \ldots, v_d$ of u so that the rotor at u points to v_{i+1} immediately after pointing to v_i (indices mod d). We assign *weight* $w(u, \eta(u)) = 0$ to a rotor pointing from u to $\eta(u)$, and weight $w(u, v_i) = H(u) - H(v_i) + w(u, v_{i-1})$ to a rotor pointing from u to v_i. These assignments are consistent since H is a harmonic function: $\sum_i (H(u) - H(v_i)) = 0$. We also assign weight $H(u)$ to a particle located at u. The sum of rotor and particle weights in any configuration is invariant under the operation of routing a particle and rotating the corresponding rotor. Initially, the sum of all particle weights is $H_w(s, Y)$. After all particles
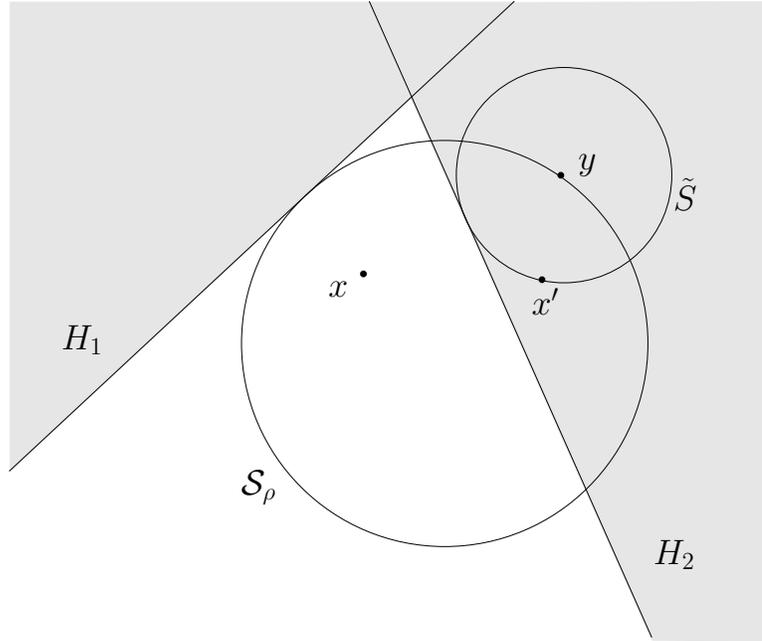
Figure 2.3: Diagram for the proof of Lemma 2.4.8.

have stopped, the sum of the particle weights is $H_r(s, Y)$. Their difference is thus at most the change in rotor weights, which is bounded above by the sum in (2.29). $\qquad \square$

For $\rho \in \mathbb{Z}$ let

$$\mathcal{S}_\rho = \{x \in \mathbb{Z}^d \ : \ \rho \le |x| < \rho + 1\}. \tag{2.30}$$

Then

$$B_\rho = \{x \in \mathbb{Z}^d \ : \ |x| < \rho\} = \mathcal{S}_0 \cup \ldots \cup \mathcal{S}_{\rho-1}.$$

Note that for simple random walk started in B_ρ, the first exit time of B_ρ and first hitting time of $\mathcal{S}_\rho$ coincide. Our next result is a modification of Lemma 5(b) of [29].

Lemma 2.4.8. *Fix $\rho \ge 1$ and $y \in \mathcal{S}_\rho$. For $x \in B_\rho$ let $H(x) = \mathbb{P}_x(X_T = y)$, where T is the first hitting time of $\mathcal{S}_\rho$. Then*

$$H(x) \le \frac{J}{|x - y|^{d-1}} \tag{2.31}$$

for a constant J depending only on d.

Proof. We induct on the distance $|x - y|$, assuming the result holds for all x' with $|x' - y| \le \frac{1}{2}|x - y|$; the base case can be made trivial by choosing J sufficiently large. By Lemma 5(b) of

[29], we can choose J large enough so that the result holds provided $|y| - |x| \geq 2^{-d-3}|x - y|$. Otherwise, let H_1 be the outer half-space tangent to $\mathcal{S}_\rho$ at the point of $\mathcal{S}_\rho$ closest to x, and let H_2 be the inner half-space tangent to the ball $\widetilde{S}$ of radius $\frac{1}{2}|x - y|$ about y, at the point of $\widetilde{S}$ closest to x. By Lemma 2.4.4 applied to these half-spaces, the probability that random walk started at x reaches $\widetilde{S}$ before hitting $\mathcal{S}_\rho$ is at most 2^{1-d}. Writing $\widetilde{T}$ for the first hitting time of $\widetilde{S} \cup \mathcal{S}_\rho$, we have

$$H(x) \leq \sum_{x' \in \widetilde{S}} \mathbb{P}_x(X_{\widetilde{T}} = x')H(x') \leq 2^{1-d} J \cdot \left(\frac{|x - y|}{2}\right)^{1-d}$$

where we have used the inductive hypothesis to bound $H(x')$. $\qquad\square$

The *lazy* random walk in $\mathbb{Z}^d$ stays in place with probability $\frac{1}{2}$, and moves to each of the $2d$ neighbors with probability $\frac{1}{4d}$. We will need the following standard result, which can be derived e.g. from the estimates in [36], section II.12; we include a proof for the sake of completeness.

Lemma 2.4.9. *Given $u \sim v \in \mathbb{Z}^d$, lazy random walks started at u and v can be coupled with probability $1 - C/R$ before either reaches distance R from u, where C depends only on d.*

Proof. Let i be the coordinate such that $u_i \neq v_i$. To define a step of the coupling, choose one of the d coordinates uniformly at random. If the chosen coordinate is different from i, let the two walks take the same lazy step so that they still agree in this coordinate. If the chosen coordinate is i, let one walk take a step while the other stays in place. With probability $\frac{1}{2}$ the walks will then be coupled. Otherwise, they are located at points u', v' with $|u' - v'| = 2$. Moreover, $\mathbb{P}\left(|u - u'| \geq \frac{R}{2\sqrt{d}}\right) < \frac{C'}{R}$ for a constant C' depending only on d. From now on, whenever coordinate i is chosen, let the two walks take lazy steps in opposite directions.

Let

$$H_1 = \left\{x \,\middle|\, x_i = \frac{u'_i + v'_i}{2}\right\}$$

be the hyperplane bisecting the segment $[u', v']$. Since the steps of one walk are reflections in H_1 of the steps of the other, the walks couple when they hit H_1. Let Q be the cube of side length $R/\sqrt{d} + 2$ centered at u, and let H_2 be a hyperplane defining one of the faces of Q. By Lemma 2.4.4 with $h_1 = 1$ and $h_2 = R/4\sqrt{d}$, the probability that one of the walks exits Q before the walks couple is at most $2d \cdot \frac{5}{2} \frac{h_1 + 1}{h_2} \left(1 + \frac{1}{2h_2}\right)^2 \leq 40 \, d^{3/2} \left(1 + 2\sqrt{d}\right)^2 / R$. $\qquad\square$

Lemma 2.4.10. *With H defined as in Lemma 2.4.8, we have*

$$\sum_{u \in B_\rho} \sum_{v \sim u} |H(u) - H(v)| \leq J' \log \rho$$

for a constant J' depending only on d.

Proof. Given $u \in B_\rho$ and $v \sim u$, by Lemma 2.4.9, lazy random walks started at u and v can be coupled with probability $1 - 2C/|u - y|$ before either reaches distance $|u - y|/2$ from u. If the walks reach this distance without coupling, by Lemma 2.4.8 each has still has probability at most $J/|u - y|^{d-1}$ of exiting B_ρ at y. By the strong Markov property it follows that

$$|H(u) - H(v)| \leq \frac{2CJ}{|u - y|^d}.$$

Summing in spherical shells about y, we obtain

$$\sum_{u \in B_\rho} \sum_{v \sim u} |H(u) - H(v)| \leq \sum_{t=1}^{\rho} d\omega_d t^{d-1} \frac{2CJ}{t^d} \leq J' \log \rho. \quad \square$$

We remark that Lemma 2.4.10 could also be inferred from Lemma 2.4.8 using [28, Thm. 1.7.1] in a ball of radius $|u - y|/2$ about u.

Proof of Theorem 1.2.1, Outer Estimate. Fix integers $\rho \geq r$, $h \geq 1$. In the setting of Proposition 2.4.7, let G be the lattice ball $B_{\rho+h+1}$, and let $Z = \mathcal{S}_{\rho+h}$. Fix $y \in \mathcal{S}_{\rho+h}$ and let $Y = \{y\}$. For $x \in \mathcal{S}_\rho$ let $s(x)$ be the number of particles stopped at x if all particles in rotor-router aggregation are stopped upon reaching $\mathcal{S}_\rho$. Write

$$H(x) = \mathbb{P}_x(X_T = y)$$

where T is the first hitting time of $\mathcal{S}_{\rho+h}$. By Lemma 2.4.8 we have

$$H_w(s, y) = \sum_{x \in \mathcal{S}_\rho} s(x) H(x) \leq \frac{JN_\rho}{h^{d-1}} \tag{2.32}$$

where

$$N_\rho = \sum_{x \in \mathcal{S}_\rho} s(x)$$

is the number of particles that ever visit the shell $\mathcal{S}_\rho$ in the course of rotor-router aggregation.

By Lemma 2.4.10 the sum in (2.29) is at most $J' \log h$, hence from Propositon 2.4.7 and (2.32) we have

$$H_r(s,y) \leq \frac{JN_\rho}{h^{d-1}} + J' \log h. \tag{2.33}$$

Let $\rho(0) = r$, and define $\rho(i)$ inductively by

$$\rho(i+1) = \min \left\{ \rho(i) + N_{\rho(i)}^{2/(2d-1)}, \ \min\{\rho > \rho(i) | N_\rho \leq N_{\rho(i)}/2\} \right\}. \tag{2.34}$$

Fixing $h < \rho(i+1) - \rho(i)$, we have

$$h^{d-1} \log h \leq N_{\rho(i)}^{\frac{2d-2}{2d-1}} \log N_{\rho(i)} \leq N_{\rho(i)};$$

so (2.33) with $\rho = \rho(i)$ simplifies to

$$H_r(s,y) \leq \frac{CN_{\rho(i)}}{h^{d-1}} \tag{2.35}$$

where $C = J + J'$.

Since all particles that visit $\mathcal{S}_{\rho(i)+h}$ during rotor-router aggregation must pass through $\mathcal{S}_{\rho(i)}$, we have

$$N_{\rho(i)+h} \leq \sum_{y \in \mathcal{S}_{\rho(i)+h}} H_r(s,y). \tag{2.36}$$

Let $M_k = \#(A_n \cap \mathcal{S}_k)$. There are at most $M_{\rho(i)+h}$ nonzero terms in the sum on the right side of (2.36), and each term is bounded above by (2.35), hence

$$M_{\rho(i)+h} \geq N_{\rho(i)+h} \frac{h^{d-1}}{CN_{\rho(i)}} \geq \frac{h^{d-1}}{2C}$$

where the second inequality follows from $N_{\rho(i)+h} \geq N_{\rho(i)}/2$. Summing over h, we obtain

$$\sum_{\rho=\rho(i)+1}^{\rho(i+1)-1} M_\rho \geq \frac{1}{2dC}(\rho(i+1) - \rho(i) - 1)^d. \tag{2.37}$$

The left side is at most $N_{\rho(i)}$, hence

$$\rho(i+1) - \rho(i) \leq (2dCN_{\rho(i)})^{1/d} \leq N_{\rho(i)}^{2/(2d-1)}$$

provided $N_{\rho(i)} \geq C' := (2dC)^{2d-1}$. Thus the minimum in (2.34) is not attained by its first argument. It follows that $N_{\rho(i+1)} \leq N_{\rho(i)}/2$, hence $N_{\rho(a \log r)} < C'$ for a sufficiently large constant a.

By the inner estimate, since the ball $B_{r-c\log r}$ is entirely occupied, we have

$$\sum_{\rho \geq r} M_\rho \leq \omega_d r^d - \omega_d (r - c\log r)^d$$

$$\leq cd\omega_d r^{d-1} \log r.$$

Write $x_i = \rho(i+1) - \rho(i) - 1$; by (2.37) we have

$$\sum_{i=0}^{a\log r} x_i^d \leq cd\omega_d r^{d-1} \log r,$$

By Jensen's inequality, subject to this constraint, $\sum x_i$ is maximized when all x_i are equal, in which case $x_i \leq C'' r^{1-1/d}$ and

$$\rho(a\log r) = r + \sum x_i \leq r + C'' r^{1-1/d} \log r. \tag{2.38}$$

Since $N_{\rho(a\log r)} < C'$ we have $N_{\rho(a\log r)+C'} = 0$; that is, no particles reach the shell $\mathcal{S}_{\rho(a\log r)+C'}$. Taking $c' = C' + C''$, we obtain from (2.38)

$$A_n \subset B_{r(1+c'r^{-1/d}\log r)}. \quad \square$$

Chapter 3

Scaling Limits for General Sources

This chapter is devoted to proving Theorems 1.1.2, 1.1.3 and 1.3.1. The proofs use many ideas from potential theory, and the relevant background is developed in section 3.1. The proof of Theorem 1.1.2 is broken into three sections, one for each of the three aggregation models. Of the three models the divisible sandpile is the most straightforward and is treated in section 3.2. The rotor-router model and internal DLA are treated in sections 3.3 and 3.4, respectively. Finally, in section 3.5 we deduce Theorem 1.3.1 for multiple point sources from Theorem 1.1.3 along with our results for single point sources.

3.1 Potential Theory Background

In this section we review the basic properties of superharmonic potentials and of the least superharmonic majorant. For more background on potential theory in general, we refer the reader to [3, 16]; for the obstacle problem in particular, see [7, 19].

3.1.1 Least Superharmonic Majorant

Since we will often be working with functions on $\mathbb{R}^d$ which may not be twice differentiable, it is desirable to define superharmonicity without making reference to the Laplacian. Instead we use the mean value property. A function u on an open set $\Omega \subset \mathbb{R}^d$ is *superharmonic* if it is lower-semicontinuous and for any ball $B(x, r) \subset \Omega$

$$u(x) \geq A_r u(x) := \frac{1}{\omega_d r^d} \int_{B(x,r)} u(y) dy. \tag{3.1}$$

Here ω_d is the volume of the unit ball in $\mathbb{R}^d$. We say that u is subharmonic if $-u$ is superharmonic, and harmonic if it is both super- and subharmonic.

The following properties of superharmonic functions are well known; for proofs, see e.g. [3], [16] or [35].

Lemma 3.1.1. *Let u be a superharmonic function on an open set $\Omega \subset \mathbb{R}^d$. Then*

(i) *u attains its minimum in $\bar{\Omega}$ on the boundary.*

(ii) *If h is harmonic on Ω and $h = u$ on $\partial\Omega$, then $u \geq h$.*

(iii) *If $B(x, r_0) \subset B(x, r_1) \subset \Omega$, then*

$$A_{r_0} u(x) \geq A_{r_1} u(x).$$

(iv) *If u is twice differentiable on Ω, then $\Delta u \leq 0$ on Ω.*

(v) *If $B \subset \Omega$ is an open ball, and v is a function on Ω which is harmonic on B, continuous on $\bar{B}$, and agrees with u on B^c, then v is superharmonic.*

Given a function γ on $\mathbb{R}^d$ which is bounded above, the *least superharmonic majorant* of γ (also called the solution to the obstacle problem with obstacle γ) is the function

$$s(x) = \inf\{f(x) | f \text{ is continuous, superharmonic and } f \geq \gamma\}. \tag{3.2}$$

Note that since γ is bounded above, the infimum is taken over a nonempty set.

Lemma 3.1.2. *Let γ be a uniformly continuous function which is bounded above, and let s be given by (3.2). Then*

(i) *s is superharmonic.*

(ii) *s is continuous.*

(iii) *s is harmonic on the domain*

$$D = \{x \in \mathbb{R}^d | s(x) > \gamma(x)\}.$$

Proof. (i) Let $f \geq \gamma$ be continuous and superharmonic. Then $f \geq s$. By the mean value property (3.1), we have

$$f \geq A_r f \geq A_r s.$$

Taking the infimum over f on the left side, we conclude that $s \geq A_r s$.

It remains to show that s is lower-semicontinuous. Let

$$\omega(\gamma, r) = \sup_{x,y \in \mathbb{R}^d, |x-y| \leq r} |\gamma(x) - \gamma(y)|.$$

Since

$$A_r s \geq A_r \gamma \geq \gamma - \omega(\gamma, r)$$

the function $A_r s + \omega(\gamma, r)$ is continuous, superharmonic, and lies above γ, so

$$A_r s \leq s \leq A_r s + \omega(\gamma, r).$$

Since γ is uniformly continuous, we have $\omega(\gamma, r) \downarrow 0$ as $r \downarrow 0$, hence $A_r s \to s$ as $r \downarrow 0$. Moreover if $r_0 < r_1$, then by Lemma 3.1.1(iii)

$$A_{r_0} s = \lim_{r \to 0} A_{r_0} A_r s \geq \lim_{r \to 0} A_{r_1} A_r s = A_{r_1} s.$$

Thus s is an increasing limit of continuous functions and hence lower-semicontinuous.

(ii) Since s is defined as an infimum of continuous functions, it is also upper-semicontinuous.

(iii) Given $x \in D$, write $\epsilon = s(x) - \gamma(x)$. Choose δ small enough so that for all $y \in B = B(x, \delta)$

$$|\gamma(x) - \gamma(y)| < \frac{\epsilon}{2} \quad \text{and} \quad |s(x) - s(y)| < \frac{\epsilon}{2}.$$

Let f be the continuous function which is harmonic in B and agrees with s outside B. By Lemma 3.1.1(v), f is superharmonic. By Lemma 3.1.1(i), f attains its minimum in $\overline{B}$ at a point $z \in \partial B$, hence for $y \in B$

$$f(y) \geq f(z) = s(z) \geq s(x) - \frac{\epsilon}{2} = \gamma(x) + \frac{\epsilon}{2} > \gamma(y).$$

It follows that $f \geq \gamma$ everywhere, hence $f \geq s$. From Lemma 3.1.1(ii) we conclude that $f = s$, and hence s is harmonic at x. $\qquad \square$

3.1.2 Superharmonic Potentials

Next we describe the particular class of obstacles which relate to the aggregation models we are studying. For a bounded measurable function σ on $\mathbb{R}^d$ with compact support, write

$$G\sigma(x) = \int_{\mathbb{R}^d} g(x, y)\sigma(y)dy, \tag{3.3}$$

where

$$g(x,y) = \begin{cases} -\frac{2}{\pi} \log |x - y|, & d = 2; \\ a_d |x - y|^{2-d}, & d \geq 3. \end{cases} \tag{3.4}$$

Here $a_d = \frac{2}{(d-2)\omega_d}$, where ω_d is the volume of the unit ball in $\mathbb{R}^d$. Note that (3.4) differs by a factor of $2d$ from the standard harmonic potential in $\mathbb{R}^d$; however, the normalization we have chosen is most convenient when working with the discrete Laplacian and random walk.

The following result is standard; see [16, Theorem 1.I.7.2].

Lemma 3.1.3. *Let σ be a measurable function on $\mathbb{R}^d$ with compact support.*

(i) *If σ is bounded, then $G\sigma$ is continuous.*

(ii) *If σ is C^1, then $G\sigma$ is C^2 and*

$$\Delta G\sigma = -2d\sigma. \tag{3.5}$$

Regarding (ii), if we remove the smoothness assumption on σ, equation (3.5) remains true in the sense of distributions. For our applications, however, we will not need this fact, and the following lemma will suffice.

Lemma 3.1.4. *Let σ be a bounded measurable function on $\mathbb{R}^d$ with compact support. If $\sigma \geq 0$ on an open set $\Omega \subset \mathbb{R}^d$, then $G\sigma$ is superharmonic on Ω.*

Proof. Suppose $B(x,r) \subset \Omega$. Since for any fixed y the function $f(x) = g(x,y)$ is superharmonic in x, we have

$$\begin{aligned} G\sigma(x) &= \int_{\mathbb{R}^d} \sigma(y) g(x,y) dy \\ &\geq \int_{\mathbb{R}^d} \sigma(y) A_r f(y) dy \\ &= A_r G\sigma(x). \quad \square \end{aligned}$$

By applying Lemma 3.1.4 both to σ and to $-\sigma$, we obtain that $G\sigma$ is harmonic off the support of σ.

Let $B = B(o, r)$ be the ball of radius r centered at the origin in $\mathbb{R}^d$. We compute in dimensions $d \geq 3$

$$G1_B(x) = \begin{cases} \frac{dr^2}{d-2} - |x|^2, & |x| < r \\ \frac{2r^2}{d-2}\left(\frac{r}{|x|}\right)^{d-2}, & |x| \geq r. \end{cases} \tag{3.6}$$

Likewise in two dimensions

$$G1_B(x) = \begin{cases} r^2(1 - 2\log r) - |x|^2, & |x| < r \\ -2r^2 \log|x|, & |x| \geq r. \end{cases} \tag{3.7}$$

Fix a bounded nonnegative function σ on $\mathbb{R}^d$ with compact support, and let

$$\gamma(x) = -|x|^2 - G\sigma(x). \tag{3.8}$$

Let

$$s(x) = \inf\{f(x) | f \text{ is continuous, superharmonic and } f \geq \gamma\} \tag{3.9}$$

be the least superharmonic majorant of γ, and let

$$D = \{x \in \mathbb{R}^d | s(x) > \gamma(x)\} \tag{3.10}$$

be the noncoincidence set.

Lemma 3.1.5. *(i) $\gamma(x) + |x|^2$ is subharmonic on $\mathbb{R}^d$.*

(ii) If $\sigma \leq M$ on an open set $\Omega \subset \mathbb{R}^d$, then $\gamma(x) - (M-1)|x|^2$ is superharmonic on Ω.

Proof. (i) By Lemma 3.1.4, since σ is nonnegative, the function $\gamma(x) + |x|^2 = -G\sigma(x)$ is subharmonic on $\mathbb{R}^d$.

(ii) Let $B = B(o, R)$ be a ball containing the support of σ. By (3.6) and (3.7), for $x \in B$ we have

$$|x|^2 = c_d R^2 - G1_B(x)$$

where $c_2 = 1 - 2\log R$ and $c_d = \frac{d}{d-2}$ for $d \geq 3$. Hence for $x \in B$ we have

$$\begin{aligned} \gamma(x) - (M-1)|x|^2 &= -G\sigma(x) - M|x|^2 \\ &= G(M1_B - \sigma)(x) - c_d M R^2. \end{aligned}$$

Since $\sigma \leq M1_B$ on Ω, by Lemma 3.1.4 the function $\gamma - (M-1)|x|^2$ is superharmonic in $B \cap \Omega$. Since this holds for all sufficiently large R, it follows that $\gamma - (M-1)|x|^2$ is superharmonic on all of Ω. $\qquad\square$

Lemma 3.1.6. *Let $u = s - \gamma$, where γ and s are given by (3.8) and (3.9). Then*

(i) $u(x) - |x|^2$ is superharmonic on $\mathbb{R}^d$.

(ii) If $\sigma \leq M$ on an open set $\Omega \subset \mathbb{R}^d$, then $u(x) + M|x|^2$ is subharmonic on Ω.

Proof. (i) By Lemmas 3.1.2(i) and 3.1.5(i), the function

$$u - |x|^2 = s - (\gamma + |x|^2)$$

is the difference of a superharmonic and a subharmonic function, hence superharmonic on $\mathbb{R}^d$.

(ii) With A_r defined by (3.1), we have

$$A_r |x|^2 = \frac{1}{\omega_d r^d} \int_{B(o,r)} (|x|^2 + 2x \cdot y + |y|^2) \, dy$$
$$= |x|^2 + \frac{1}{\omega_d r^d} \int_0^r (d\omega_d t^{d-1}) t^2 \, dt$$
$$= |x|^2 + \frac{dr^2}{d+2}.$$

By Lemma 3.1.5 we have

$$A_r s + \frac{dr^2}{d+2} \geq A_r \gamma + A_r |x|^2 - |x|^2 \geq \gamma.$$

Since $A_r s$ is continuous and superharmonic, it follows that $A_r s + \frac{dr^2}{d+2} \geq s$, hence

$$A_r s + A_r |x|^2 = A_r s + |x|^2 + \frac{dr^2}{d+2} \geq s + |x|^2.$$

Thus $s + |x|^2$ is subharmonic on $\mathbb{R}^d$, and hence by Lemma 3.1.5(ii) the function

$$u + M|x|^2 = (s + |x|^2) - (\gamma - (M-1)|x|^2)$$

is subharmonic on Ω. $\square$

For $A \subset \mathbb{R}^d$, write A^o for the interior of A and $\bar{A}$ for the closure of A. The boundary of A is $\partial A = \bar{A} - A^o$.

Lemma 3.1.7. *Let D be given by (3.10). Then $\{\sigma > 1\}^o \subset D$.*

Proof. If $\sigma > 1$ in a ball $B = B(x, r)$, by (3.6) and (3.7), for $y \in B$ we have

$$\gamma(y) = -|y|^2 - G\sigma(y) = -c_d r^2 - G(\sigma - 1_B)(y).$$

By Lemma 3.1.4 it follows that γ is subharmonic in B. In particular, $s > \gamma$ in B, so $x \in D$. $\qquad\square$

The next lemma concerns the monotonicity of our model: by starting with more mass, we obtain a larger odometer and a larger noncoincidence set; see also [41].

Lemma 3.1.8. *Let $\sigma_1 \leq \sigma_2$ be functions on $\mathbb{R}^d$ with compact support, and suppose that $\int_{\mathbb{R}^d} \sigma_2(x)dx < \infty$. Let*

$$\gamma_i(x) = -|x|^2 - G\sigma_i(x), \qquad i = 1, 2.$$

Let s_i be the least superharmonic majorant of γ_i, let $u_i = s_i - \gamma_i$, and let

$$D_i = \{x | s_i(x) > \gamma_i(x)\}.$$

Then $u_1 \leq u_2$ and $D_1 \subset D_2$.

Proof. Let

$$\tilde{s} = s_2 + G(\sigma_2 - \sigma_1).$$

Then $\tilde{s}$ is continuous and superharmonic, and since $s_2(x) \geq -|x|^2 - G\sigma_2(x)$ we have

$$\tilde{s}(x) \geq -|x|^2 - G\sigma_1(x)$$

hence $\tilde{s} \geq s_1$. Now

$$u_2 - u_1 = s_2 - s_1 + G(\sigma_2 - \sigma_1) = \tilde{s} - s_1 \geq 0.$$

Since D_i is the support of u_i, the result follows. $\qquad\square$

Our next lemma shows that we can restrict to a domain $\Omega \subset \mathbb{R}^d$ when taking the least superharmonic majorant, provided that Ω contains the noncoincidence set.

Lemma 3.1.9. *Let γ, s, D be given by (3.8)-(3.10). Let $\Omega \subset \mathbb{R}^d$ be an open set with $D \subset \Omega$. Then*

$$s(x) = \inf\{f(x) | f \text{ is superharmonic on } \Omega, \text{ continuous, and } f \geq \gamma\}. \qquad (3.11)$$

Proof. Let f be any continuous function which is superharmonic on Ω and $\geq \gamma$. By Lemma 3.1.2(iii), s is harmonic on D, so $f - s$ is superharmonic on D and attains its minimum in $\overline{D}$ on the boundary. Hence $f - s$ attains its minimum in $\overline{\Omega}$ at a point x where $s(x) = \gamma(x)$. Since $f \geq \gamma$ we conclude that $f \geq s$ on Ω and hence everywhere. Thus s is at most the infimum in (3.11). Since the infimum in (3.11) is taken over a strictly larger set than that in (3.9), the reverse inequality is trivial. $\square$

3.1.3   Boundary Regularity for the Obstacle Problem

Next we turn to the regularity of the solution to the obstacle problem (3.9) and of the free boundary ∂D. There is a substantial literature on boundary regularity for the obstacle problem. In our setup, however, extra care must be taken near points where $\sigma(x) = 1$: at these points the obstacle (3.8) is harmonic, and the free boundary can be badly behaved. We show only the minimal amount of regularity required for the proofs of our main theorems. Much stronger regularity results are known in related settings; see, for example, [7, 8].

The following lemma shows that if the obstacle is sufficiently smooth, then the superharmonic majorant cannot separate too quickly from the obstacle near the boundary of the noncoincidence set. For the proof, we follow the sketch in Caffarelli [7, Theorem 2]. As usual, we write D_ϵ for the inner ϵ-neighborhood of D, given by (1.8)

Lemma 3.1.10. *Let σ be a C^1 function on $\mathbb{R}^d$ with compact support. Let γ, s, D be given by (3.8)-(3.10), and write $u = s - \gamma$. Then u is C^1, and for $y \in \partial D_\epsilon$ we have $|\nabla u(y)| \leq C_0 \epsilon$, for a constant C_0 depending on σ.*

Proof. Fix $x_0 \in \partial D$, and define

$$L(x) = \gamma(x_0) + \langle \nabla \gamma(x_0), x - x_0 \rangle.$$

Since σ is C^1, we have that γ is C^2 by Lemma 3.1.3(ii). Let A be the maximum second partial of γ in the ball $B = B(x_0, 4\epsilon)$. By the mean value theorem and Cauchy-Schwarz, for $x \in B$ we have

$$\begin{aligned}
|L(x) - \gamma(x)| &= |\langle \nabla \gamma(x_0) - \nabla \gamma(x_*), x - x_0 \rangle| \\
&\leq A\sqrt{d}|x_0 - x_*||x - x_0| \leq c\epsilon^2,
\end{aligned} \tag{3.12}$$

where $c = 16A\sqrt{d}$. Hence

$$s(x) \geq \gamma(x) \geq L(x) - c\epsilon^2, \qquad x \in B.$$

Thus the function

$$w = s - L + c\epsilon^2$$

is nonnegative and superharmonic in B. Write $w = w_0 + w_1$, where w_0 is harmonic and equal to w on ∂B. Then since $s(x_0) = \gamma(x_0)$, we have

$$w_0(x_0) \leq w(x_0) = s(x_0) - L(x_0) + c\epsilon^2 = c\epsilon^2.$$

By the Harnack inequality, it follows that $w_0 \leq c'\epsilon^2$ on the ball $B' = B(x_0, 2\epsilon)$, for a suitable constant c'.

Since w_1 is nonnegative and vanishes on ∂B, it attains its maximum in $\bar{B}$ at a point x_1 in the support of its Laplacian. Since $\Delta w_1 = \Delta s$, by Lemma 3.1.2(iii) we have $s(x_1) = \gamma(x_1)$, hence

$$w_1(x_1) \leq w(x_1) = s(x_1) - L(x_1) + c\epsilon^2 \leq 2c\epsilon^2,$$

where in the last step we have used (3.12). We conclude that $0 \leq w \leq (2c + c')\epsilon^2$ on B' and hence $|s - L| \leq (c + c')\epsilon^2$ on B'. Thus on B' we have

$$|u| = |s - \gamma| \leq |s - L| + |\gamma - L| \leq (2c + c')\epsilon^2.$$

In particular, u is differentiable at x_0, and $\nabla u(x_0) = 0$. Since s is harmonic in D and equal to γ outside D, it follows that u is differentiable everywhere, and C^1 off ∂D.

Given $y \in \partial D_\epsilon$, let x_0 be the closest point in ∂D to y. Since $B(y, \epsilon) \subset D$, by Lemma 3.1.2(iii) the function w is harmonic on $B(y, \epsilon)$, so by the Cauchy estimate [3, Theorem 2.4] we have

$$|\nabla s(y) - \nabla \gamma(x_0)| = |\nabla w(y)| \leq \frac{C}{\epsilon} \sup_{z \in B(y,\epsilon)} w(z). \tag{3.13}$$

Since $B(y, \epsilon) \subset B'$, the right side of (3.13) is at most $C(2c + c')\epsilon$, hence

$$|\nabla u(y)| \leq |\nabla s(y) - \nabla \gamma(x_0)| + |\nabla \gamma(y) - \nabla \gamma(x_0)| \leq C_0\epsilon$$

where $C_0 = C(c + c') + A\sqrt{d}$. Thus u is C^1 on ∂D as well. $\qquad\square$

Next we show that mass is conserved in our model: the amount of mass starting in D is $\int_D \sigma(x)dx$, while the amount of mass ending in D is $\mathcal{L}(D)$, the Lebesgue measure of D. Since no mass moves outside of D, we expect these to be equal. Although this seems intuitively obvious, the proof takes some work because we have no *a priori* control of the domain D. In particular, we first need to show that the boundary of D cannot have positive d-dimensional Lebesgue measure.

Proposition 3.1.11. *Let σ be a C^1 function on $\mathbb{R}^d$ with compact support, such that $\mathcal{L}(\sigma^{-1}(1)) = 0$. Let D be given by (3.10). Then*

(i) $\mathcal{L}(\partial D) = 0$.

(ii) For any function $h \in C^1(\bar{D})$ which is superharmonic on D,

$$\int_D h(x)dx \leq \int_D h(x)\sigma(x)dx.$$

Note that by applying (ii) both to h and to $-h$, equality holds whenever h is harmonic on D. In particular, taking $h = 1$ yields the conservation of mass: $\int_D \sigma(x)dx = \mathcal{L}(D)$.

The proof of part (i) follows Friedman [19, Ch. 2, Theorem 3.5]. It uses the Lebesgue density theorem, as stated in the next lemma.

Lemma 3.1.12. *Let $A \subset \mathbb{R}^d$ be a Lebesgue measurable set, and let*

$$A' = \big\{ x \in A \,|\, \liminf_{\epsilon \to 0} \epsilon^{-d} \mathcal{L}\big(B(x,\epsilon) \cap A^c\big) > 0 \big\}.$$

Then $\mathcal{L}(A') = 0$.

To prove the first part of Proposition 3.1.11, given a boundary point $x \in \partial D$, the idea is first to find a point y in the ball $B(x, \epsilon)$ where $u = s - \gamma$ is relatively large, and then to argue using Lemma 3.1.10 that a ball $B(y, c\epsilon)$ must be entirely contained in D. Taking $A = \partial D$ in the Lebesgue density theorem, we obtain that $x \in A'$.

The proof of the second part of Proposition 3.1.11 uses Green's theorem in the form

$$\int_{D'} (u\Delta h - \Delta u h)dx = \int_{\partial D'} \left(u\frac{\partial h}{\partial \mathbf{n}} - \frac{\partial u}{\partial \mathbf{n}}h \right) dr. \tag{3.14}$$

Here D' is the union of boxes $x^\square$ that are contained in D, and dx is the volume measure in D', while dr is the $(d-1)$-dimensional surface measure on $\partial D'$. The partial derivatives on the right side of (3.14) are in the outward normal direction from $\partial D'$.

Proof of Propostion 3.1.11. (i) Fix $0 < \lambda < 1$. For $x \in \partial D$ with $\sigma(x) \leq \lambda$, for small enough ϵ we have $\sigma \leq \frac{1+\lambda}{2}$ on $B(x, \epsilon)$. By Lemma 3.1.5(ii) the function

$$f(y) = \gamma(y) + \frac{1-\lambda}{2}|y|^2$$

is superharmonic on $B(x, \epsilon) \cap D$, so by Lemma 3.1.2(iii) the function

$$w(y) = u(y) - \frac{1-\lambda}{2}|x - y|^2$$
$$= s(y) - f(y) + (1 - \lambda)\langle x, y\rangle - \frac{1-\lambda}{2}|x|^2$$

is subharmonic on $B(x, \epsilon) \cap D$. Since $w(x) = 0$, its maximum is not attained on ∂D, so it must be attained on $\partial B(x, \epsilon)$, so there is a point y with $|x - y| = \epsilon$ and

$$u(y) \geq \frac{1-\lambda}{2}\epsilon^2.$$

By Lemma 3.1.10 we have $|\nabla u| \leq (1+c)C_0\epsilon$ in the ball $B(y, c\epsilon)$. Taking $c = \frac{1-\lambda}{4C_0}$, we obtain for $z \in B(y, c\epsilon)$

$$u(z) = u(y) + \langle \nabla u(y_*), z - y\rangle$$
$$\geq \frac{1-\lambda}{2}\epsilon^2 - c(1+c)C_0\epsilon^2 > 0.$$

Thus for any $x \in \partial D \cap \{\sigma \leq \lambda\}$

$$\mathcal{L}(B(x, (1+c)\epsilon) \cap (\partial D)^c) \geq \omega_d(c\epsilon)^d.$$

By the Lebesgue density theorem it follows that

$$\mathcal{L}(\partial D \cap \{\sigma \leq \lambda\}) = 0. \tag{3.15}$$

By Lemma 3.1.7 we have $\sigma \leq 1$ on ∂D. Taking $\lambda \uparrow 1$ in (3.15), we obtain

$$\mathcal{L}(\partial D) \leq \mathcal{L}(\partial D \cap \{\sigma < 1\}) + \mathcal{L}(\sigma^{-1}(1)) = 0.$$

(ii) Fix $\delta > 0$ and let

$$D' = \bigcup_{x \in \delta\mathbb{Z}^d \,:\, x^\square \subset D} x^\square,$$

where $x^\square = x + [-\frac{\delta}{2}, \frac{\delta}{2}]^d$. By Lemmas 3.1.2(iii) and 3.1.3(ii), in D' we have

$$\Delta u = \Delta s - \Delta \gamma = 2d(1 - \sigma).$$

Now by Green's theorem (3.14), since u is nonnegative and h is superharmonic,

$$\int_{D'} (1-\sigma)h \, dx = \frac{1}{2d} \int_{D'} \Delta u h \, dx \le \frac{1}{2d} \int_{\partial D'} \left(u \frac{\partial h}{\partial \mathbf{n}} - \frac{\partial u}{\partial \mathbf{n}} h \right) dr \tag{3.16}$$

where $\mathbf{n}$ denotes the unit outward normal vector to $\partial D'$. By Lemma 3.1.10, the integral on the right side is bounded by

$$\int_{\partial D'} \left| u \frac{\partial h}{\partial \mathbf{n}} \right| + \left| \frac{\partial u}{\partial \mathbf{n}} h \right| dr \le C_0 \sqrt{d} (\delta^2 ||\nabla h||_\infty + \delta ||h||_\infty) \mathcal{L}_{d-1}(\partial D'), \tag{3.17}$$

where $\mathcal{L}_{d-1}$ denotes $(d-1)$-dimensional Lebesgue measure. Let

$$S = \bigcup_{x \in \delta \mathbb{Z}^d \, : \, x^\square \cap \partial D \ne \emptyset} x^\square.$$

Since $\mathcal{L}(\partial D) = 0$, given $\epsilon > 0$ we can choose δ small enough so that $\mathcal{L}(S) < \epsilon$. Since $\partial D' \subset \partial S$, we have

$$\mathcal{L}_{d-1}(\partial D') \le \mathcal{L}_{d-1}(\partial S) \le \frac{2d\mathcal{L}(S)}{\delta}.$$

Since D is open, $\mathcal{L}(D') \uparrow \mathcal{L}(D)$ as $\delta \downarrow 0$. Taking $\delta < \epsilon$ smaller if necessary so that $\mathcal{L}(D') \ge \mathcal{L}(D) - \epsilon$, we obtain from (3.16) and (3.17)

$$\int_D (1-\sigma)h \, dx \le (M+1)\epsilon + C_0 \sqrt{d}(||h||_\infty + ||\nabla h||_\infty)\epsilon$$

where M is the maximum of $|\sigma|$. Since this holds for any $\epsilon > 0$, we conclude that $\int_D h(x)dx \le \int_D h(x)\sigma(x)dx$. $\qquad\square$

We will need a version of Proposition 3.1.11 which does not assume that σ is C^1 or even continuous. We can replace the C^1 assumption and the condition that $\mathcal{L}(\sigma^{-1}(1)) = 0$ by the following condition.

$$\text{For all } x \in \mathbb{R}^d \text{ either } \sigma(x) \le \lambda \text{ or } \sigma(x) \ge 1 \tag{3.18}$$

for a constant $\lambda < 1$. Then we have the following result.

Proposition 3.1.13. *Let σ be a bounded function on $\mathbb{R}^d$ with compact support. Let D be given by (3.10), and let $\widetilde{D} = D \cup \{\sigma \ge 1\}^o$. If σ is continuous almost everywhere and satisfies (3.18), then*

(i) $\mathcal{L}(\partial \widetilde{D}) = 0$.

(ii) $\mathcal{L}(D) = \int_D \sigma(x)dx$.

(iii) $\mathcal{L}(\widetilde{D}) = \int_{\widetilde{D}} \sigma(x) dx.$

In particular, taking $\sigma = 1_A + 1_B$, we have

$$A \oplus B = A \cup B \cup D = \widetilde{D}$$

where $\oplus$ denotes the smash sum (1.6). From (iii) we obtain the volume additivity of the smash sum.

Corollary 3.1.14. *Let $A, B \subset \mathbb{R}^d$ be bounded open sets whose boundaries have measure zero. Then*

$$\mathcal{L}(A \oplus B) = \mathcal{L}(A) + \mathcal{L}(B).$$

Proof of Proposition 3.1.13. (i) Fix $\epsilon > 0$. Since σ is continuous almost everywhere, there exist C^1 functions $\sigma_0 \leq \sigma \leq \sigma_1$ with $\int_{\mathbb{R}^d}(\sigma_1 - \sigma_0)dx < \epsilon$. Scaling by a factor of $1 + \delta$ for sufficiently small δ, we can ensure that $\mathcal{L}(\sigma_i^{-1}(1)) = 0$, so that σ_0 and σ_1 satisfy the hypotheses of Proposition 3.1.11. For $i = 0, 1$ let

$$\gamma_i(x) = -|x|^2 - G\sigma_i(x),$$

and let s_i be the least superharmonic majorant of γ_i. Choose α with $\lambda < \alpha < 1$ such that $\mathcal{L}(\sigma_i^{-1}(\alpha)) = 0$ for $i = 0, 1$, and write

$$D_i = \{s_i > \gamma_i\};$$
$$S_i = D_i \cup \{\sigma_i \geq \alpha\}.$$

By (3.18) we have $\{\sigma_0 \geq \alpha\} \subset \{\sigma \geq 1\}^o$, hence by Lemma 3.1.8

$$S_0 = D_0 \cup \{\sigma_0 \geq \alpha\} \subset D \cup \{\sigma \geq 1\}^o = \widetilde{D} \tag{3.19}$$
$$\subset \overline{\widetilde{D}} \subset \overline{D_1 \cup \{\sigma \geq 1\}} = \overline{S_1}. \tag{3.20}$$

For $i = 0, 1$ write

$$\sigma_i^\circ(x) = \begin{cases} 1, & x \in D_i \\ \sigma_i(x), & x \notin D_i. \end{cases}$$

By Proposition 3.1.11(ii) with $h = 1$, we have

$$\int_{\mathbb{R}^d} \sigma_i^\circ = \mathcal{L}(D_i) + \int_{D_i^c} \sigma_i = \int_{\mathbb{R}^d} \sigma_i. \tag{3.21}$$

For $0 < \alpha_0 < \alpha$, we have

$$\mathcal{L}(S_1 - S_0) \leq \mathcal{L}\big(S_1 \cap \{\sigma_0^\circ \leq \alpha_0\}\big) + \mathcal{L}\big(\{\alpha_0 < \sigma_0^\circ < \alpha\}\big). \tag{3.22}$$

Since $\sigma_1^\circ \geq \alpha$ on S_1, the first term is bounded by

$$\mathcal{L}\big(S_1 \cap \{\sigma_0^\circ \leq \alpha_0\}\big) \leq \frac{||\sigma_1^\circ - \sigma_0^\circ||_1}{\alpha - \alpha_0} = \frac{||\sigma_1 - \sigma_0||_1}{\alpha - \alpha_0} < \frac{\epsilon}{\alpha - \alpha_0},$$

where the equality in the middle step follows from (3.21).

By Proposition 3.1.11(i) we have $\mathcal{L}(\partial D_1) = 0$, hence $\mathcal{L}(\partial S_1) = 0$. Taking $\alpha_0 = \alpha - \sqrt{\epsilon}$ we obtain from (3.19), (3.20), and (3.22)

$$\mathcal{L}\big(\partial \widetilde{D}\big) \leq \mathcal{L}\big(\overline{S_1} - S_0\big) = \mathcal{L}(S_1 - S_0) < \sqrt{\epsilon} + \mathcal{L}\big(\{\alpha - \sqrt{\epsilon} < \sigma_0 < \alpha\}\big).$$

Since this holds for any $\epsilon > 0$, the result follows.

(ii) Write

$$\sigma^\circ(x) = \begin{cases} 1, & x \in D \\ \sigma(x), & x \notin D. \end{cases}$$

From (3.21) we have

$$\int_{\mathbb{R}^d} \sigma_0 = \int_{\mathbb{R}^d} \sigma_0^\circ \leq \int_{\mathbb{R}^d} \sigma^\circ \leq \int_{\mathbb{R}^d} \sigma_1^\circ = \int_{\mathbb{R}^d} \sigma_1.$$

The left and right side differ by at most ϵ. Since $\int_{\mathbb{R}^d} \sigma_0 \leq \int_{\mathbb{R}^d} \sigma \leq \int_{\mathbb{R}^d} \sigma_1$, and $\epsilon > 0$ is arbitrary, it follows that $\int_{\mathbb{R}^d} \sigma = \int_{\mathbb{R}^d} \sigma^\circ$, hence

$$\int_D \sigma = \int_{\mathbb{R}^d} \sigma^\circ - \int_{D^c} \sigma = \mathcal{L}(D). \tag{3.23}$$

(iii) By Lemma 3.1.7, if $\sigma(x) > 1$ and $x \notin D$, then σ is discontinuous at x. Thus $\sigma = 1$ almost everywhere on $\widetilde{D} - D$, and we obtain from (3.23)

$$\int_{\widetilde{D}} \sigma = \mathcal{L}(D) + \int_{\widetilde{D} - D} \sigma = \mathcal{L}(\widetilde{D}). \quad \square$$

Our next lemma describes the domain resulting from starting mass $m > 1$ on a ball in $\mathbb{R}^d$. Not surprisingly, the result is another ball, concentric with the original, of m times the volume. In particular, if m is an integer, the m-fold smash sum of a ball with itself is again a ball.

Lemma 3.1.15. *Fix $m > 1$, and let D be given by (3.10) with $\sigma = m1_{B(o,r)}$. Then $D = B(o, m^{1/d}r)$.*

Proof. Since $\gamma(x) = -|x|^2 - G\sigma(x)$ is spherically symmetric and the least superharmonic majorant commutes with rotations, D is a ball centered at the origin. By Proposition 3.1.13(ii) we have $\mathcal{L}(D) = m\mathcal{L}(B(o,r))$. $\square$

Next we show that the noncoincidence set is bounded; see also [42, Cor. 7.2].

Lemma 3.1.16. *Let σ be a function on $\mathbb{R}^d$ with compact support, satisfying $0 \leq \sigma \leq M$. Let D be given by (3.10). Then D is bounded.*

Proof. Let $B = B(o, r)$ be a ball containing the support of σ. By Lemmas 3.1.8 and 3.1.15, the difference $s - \gamma$ is supported in $B(o, M^{1/d}r)$. $\square$

3.1.4 Convergence of Obstacles, Majorants and Domains

In broad outline, many of our arguments have the following basic structure:

$$\text{convergence of densities} \implies \text{convergence of obstacles}$$
$$\implies \text{convergence of majorants}$$
$$\implies \text{convergence of domains.}$$

An appropriate convergence of starting densities is built into the hypotheses of the theorems. From these hypotheses we use Green's function estimates to deduce the relevant convergence of obstacles. Next, as we have already seen in Lemmas 3.1.5 and 3.1.6, properties of the obstacle can often be parlayed into corresponding properties of the least superharmonic majorant. Finally, deducing convergence of domains (i.e., noncoincidence sets) from the convergence of the majorants often requires additional regularity assumptions. The following lemma illustrates this basic three-step approach.

Lemma 3.1.17. *Let σ and σ_n, $n = 1, 2, \ldots$ be densities on $\mathbb{R}^d$ satisfying*

$$0 \leq \sigma, \sigma_n \leq M1_B$$

for a constant M and ball $B = B(o, R)$. Suppose that σ is continuous except on a set of Lebesgue measure zero, and that

$$\sigma_n(x) \to \sigma(x) \tag{3.24}$$

as $n \to \infty$, for all continuity points x of σ. Then

(i) $G\sigma_n \to G\sigma$ uniformly on compact subsets of $\mathbb{R}^d$.

(ii) $s_n \to s$ uniformly on compact subsets of $\mathbb{R}^d$, where s, s_n are the least superharmonic majorants of the functions $\gamma = -|x|^2 - G\sigma$, and $\gamma_n = -|x|^2 - G\sigma_n$, respectively.

(iii) For any $\epsilon > 0$ we have $D_\epsilon \subset D_{(n)}$ for all sufficiently large n, where $D, D_{(n)}$ are the noncoincidence sets $\{s > \gamma\}$ and $\{s_n > \gamma_n\}$, respectively.

Proof. (i) Since σ, σ_n are supported on B, we have

$$|G\sigma_n(x) - G\sigma(x)| \leq \int_B |g(x,y)||\sigma_n(y) - \sigma(y)|dy. \tag{3.25}$$

Fix $0 < \epsilon < 1$. Since $0 \leq \sigma, \sigma_n \leq M$, we have by (3.6) and (3.7)

$$\int_{B(x,\epsilon)} |g(x,y)||\sigma_n(y) - \sigma(y)|dy \leq 3M\epsilon^2. \tag{3.26}$$

Now let $B_0 = B - B(x,\epsilon)$ and

$$A_N = \{y \in B_0 \ : \ |\sigma_n(y) - \sigma(y)| < \epsilon \text{ for all } n \geq N\}.$$

Then by (3.24)

$$\bigcup_{N \geq 1} A_N \supseteq B_0 - DC(\sigma)$$

where $DC(\sigma)$ is the set of discontinuities of σ. As $\mathcal{L}(DC(\sigma)) = 0$ and the sets A_N are monotone increasing in N, we can choose N large enough so that

$$\mathcal{L}(A_N) > \mathcal{L}(B_0) - \frac{\epsilon^{d-1}}{\log(R/\epsilon)}. \tag{3.27}$$

Given $n \geq N$, for $y \in B_0$ we have

$$|g(x,y)| \leq a_d \epsilon^{2-d} \log \frac{R}{\epsilon} \leq \frac{a_d \epsilon}{\mathcal{L}(B_0 - A_n)}.$$

Splitting the right side of (3.25) into separate integrals over $B(x,\epsilon)$, A_n, and $B_0 - A_n$, we obtain

$$|G\sigma_n(x) - G\sigma(x)| < 6M\epsilon^2 + \epsilon \int_{A_n} |g(x,y)|dy + 2Ma_d\epsilon$$

$$\leq \left(3R^2 \log\max(R,|x|) + (2a_d + 3)M\right)\epsilon.$$

(ii) By Lemmas 3.1.8 and 3.1.15, the noncoincidence sets $D, D_{(n)}$ are contained in the ball $B_1 = M^{1/d}B$. Given $\epsilon > 0$ and a compact set K containing B_1, choose N large enough so that $|\gamma_n - \gamma| < \epsilon$ on K for all $n \geq N$. Then

$$s + \epsilon \geq \gamma + \epsilon > \gamma_n$$

on K, so the function $f_n = \max(s + \epsilon, \gamma_n)$ is superharmonic on K. By Lemma 3.1.9 we have $f_n \geq s_n$, and hence $s + \epsilon \geq s_n$ on K. Likewise

$$s_n + \epsilon \geq \gamma_n + \epsilon > \gamma$$

on K, so the function $\tilde{f}_n = \max(s_n + \epsilon, \gamma)$ is superharmonic on K. By Lemma 3.1.9 we have $\tilde{f}_n \geq s$ and hence $s_n + \epsilon \geq s$ on K. Thus $s_n \to s$ uniformly on K.

(iii) Let $\beta > 0$ be the minimum value of $s - \gamma$ on $\overline{D_\epsilon}$, and choose N large enough so that $|\gamma - \gamma_n|, |s - s_n| < \beta/2$ on $\overline{D}$ for all $n \geq N$. Then

$$s_n - \gamma_n > s - \gamma - \beta > 0$$

on D_ϵ, hence $D_\epsilon \subset D_{(n)}$ for all $n \geq N$. $\qquad\square$

According to the following lemma, a nonnegative function on $\mathbb{Z}^d$ with bounded Laplacian can grow at most quadratically. We will use this fact repeatedly.

Lemma 3.1.18. *Fix $0 < \beta < 1$. There is a constant c_β such that any nonnegative function f on $\mathbb{Z}^d$ with $f(o) = 0$ and $|\Delta f| \leq \lambda$ in $B(o, R)$ satisfies*

$$f(x) \leq c_\beta \lambda |x|^2, \qquad x \in B(o, \beta R).$$

Proof. Consider the functions

$$f_\pm(y) = f(y) \pm \lambda |y|^2. \tag{3.28}$$

By hypothesis, f_+ is subharmonic and f_- is superharmonic on $B(o, R)$. Given $x \in B(o, \beta R)$, let $r = \min(R, 2|x|)$, and let $h_\pm$ be the harmonic function on the ball $B = B(o, r - 1)$ which agrees with $f_\pm$ on ∂B. Then

$$f_+ \leq h_+ \leq h_- + 2\lambda r^2 \leq f_- + 2\lambda r^2. \tag{3.29}$$

By the Harnack inequality [28, Theorem 1.7.2] there is a constant $\tilde{c}_\beta$ such that

$$h_+(x) \leq \tilde{c}_\beta h_+(o).$$

Since $f_-(o) = f(o) = 0$ we have from (3.29)

$$h_+(o) \leq f_-(o) + 2\lambda r^2 \leq 8\lambda |x|^2.$$

hence

$$f(x) \leq h_+(x) - \lambda |x|^2 \leq c_\beta \lambda |x|^2$$

with $c_\beta = 8\tilde{c}_\beta - 1$. $\qquad\square$

The following continuous version of Lemma 3.1.18 is proved in the same way, using the continuous Harnack inequality in place of the discrete one, and replacing $|y|^2$ by $|y|^2/2d$ in (3.28).

Lemma 3.1.19. *Fix* $0 < \beta < 1$. *There is a constant* c'_β *such that the following holds. Let* f *be a nonnegative function on* $\mathbb{R}^d$ *with* $f(o) = 0$, *and let*

$$f_\pm(x) = f(x) \pm \lambda \frac{|x|^2}{2d}.$$

If f_+ *is subharmonic and* f_- *is superharmonic in* $B(o, R)$, *then*

$$f(x) \leq c'_\beta \lambda |x|^2, \qquad x \in B(o, \beta R).$$

In the following lemma, A_ϵ denotes the inner ϵ-neighborhood of A, as defined by (1.8).

Lemma 3.1.20. *Let* $A, B \subset \mathbb{R}^d$ *be bounded open sets. For any* $\epsilon > 0$ *there exists* $\eta > 0$ *with*

$$(A \cup B)_\epsilon \subset A_\eta \cup B_\eta.$$

Proof. Let K be the closure of $(A \cup B)_\epsilon$, and let $Y_n = A_{1/n} \cup B_{1/n}$. Since K is contained in $\bigcup Y_n = A \cup B$, the sets $K \cap Y_n$ form an open cover of K, which has a finite subcover, i.e. $K \subset Y_n$ for some n. $\qquad\square$

3.1.5 Discrete Potential Theory

Fix a sequence $\delta_n \downarrow 0$ with $\delta_1 = 1$. In this section we relate discrete superharmonic potentials in the lattice $\delta_n \mathbb{Z}^d$ to their continuous couterparts in $\mathbb{R}^d$. If A is a domain in $\delta_n \mathbb{Z}^d$, write

$$A^\square = A + \left[-\frac{\delta_n}{2}, \frac{\delta_n}{2} \right]^d$$

	$\delta_n \mathbb{Z}^d$	$\mathbb{R}^d$
points	$x^{\because} = \left(x + \left(-\frac{\delta_n}{2}, \frac{\delta_n}{2} \right]^d \right) \cap \delta_n \mathbb{Z}^d$	$x^{\square} = x + \left[-\frac{\delta_n}{2}, \frac{\delta_n}{2} \right]^d$
sets	$A^{\because} = A \cap \delta_n \mathbb{Z}^d$	$A^{\square} = A + \left[\frac{-\delta_n}{2}, \frac{\delta_n}{2} \right]^d$
functions	$f^{\because} = f\vert_{\delta_n \mathbb{Z}^d}$	$f^{\square}(x) = f(x^{\because})$

Table 3.1: Notation for transitioning between Euclidean space and the lattice.

for the corresponding union of cubes in $\mathbb{R}^d$. If A is a domain in $\mathbb{R}^d$, write $A^{\because} = A \cap \delta_n \mathbb{Z}^d$. Given $x \in \mathbb{R}^d$, write

$$x^{\because} = \left(x - \frac{\delta_n}{2}, x + \frac{\delta_n}{2} \right]^d \cap \delta_n \mathbb{Z}^d$$

for the closest lattice point to x, breaking ties to the right. For a function f on $\delta_n \mathbb{Z}^d$, write

$$f^{\square}(x) = f(x^{\because})$$

for the corresponding step function on $\mathbb{R}^d$. Likewise, for a function f on $\mathbb{R}^d$, write $f^{\because} = f\vert_{\delta_n \mathbb{Z}^d}$. These notations are summarized in Table 1.

We define the discrete Laplacian of a function f on $\delta_n \mathbb{Z}^d$ to be

$$\Delta f(x) = \delta_n^{-2} \left(\frac{1}{2d} \sum_{y \sim x} f(y) - f(x) \right).$$

According to the following lemma, if f is sufficiently smooth, then its discrete Laplacian on $\delta_n \mathbb{Z}^d$ approximates its Laplacian on $\mathbb{R}^d$.

Lemma 3.1.21. *If f has continuous third derivative in a δ_n-neighborhood of $x \in \delta_n \mathbb{Z}^d$, and A is the maximum pure third partial of f in this neighborhood, then*

$$|\Delta f(x) - 2d \Delta f^{\because}(x)| \leq \frac{d}{3} A \delta_n.$$

Proof. By Taylor's theorem with remainder

$$f(x + \delta_n e_i) - 2f(x) + f(x - \delta_n e_i) = \frac{\partial^2 f}{\partial x_i^2} \delta_n^2 + \frac{1}{6} \frac{\partial^3 f}{\partial x_i^3}(x + te_i)\delta_n^3 - \frac{1}{6} \frac{\partial^3 f}{\partial x_i^3}(x - ue_i)\delta_n^3$$

for some $0 \leq t, u \leq 1$. Summing over $i = 1, \ldots, d$ and dividing by δ_n^2 gives the result. $\square$

In three and higher dimensions, for $x, y \in \delta_n \mathbb{Z}^d$ write

$$g_n(x, y) = \delta_n^{2-d} g_1(\delta_n^{-1} x, \delta_n^{-1} y), \qquad d \geq 3, \tag{3.30}$$

where

$$g_1(x,y) = \mathbb{E}_x \#\{k | X_k = y\}$$

is the Green's function for simple random walk on $\mathbb{Z}^d$. The scaling in (3.30) is chosen so that $\Delta_x g_n(x,y) = -\delta_n^{-d} 1_{\{x=y\}}$. In two dimensions, write

$$g_n(x,y) = -a(\delta_n^{-1}x, \delta_n^{-1}y) + \frac{2}{\pi} \log \delta_n, \qquad d = 2, \tag{3.31}$$

where

$$a(x,y) = \lim_{m \to \infty} \left(\mathbb{E}_x \#\{k \le m | X_k = x\} - \mathbb{E}_x \#\{k \le m | X_k = y\} \right)$$

is the recurrent potential kernel on $\mathbb{Z}^2$.

Lemma 3.1.22. *In all dimensions $d \ge 2$,*

$$g_n(x,y) = g(x,y) + O(\delta_n^2 |x-y|^{-d})$$

where g is given by (3.4).

Proof. In dimensions $d \ge 3$ we have from (3.30) and the standard estimate for the discrete Green's function (2.5)

$$g_n(x,y) = \delta_n^{2-d} \left(a_d \left(\frac{|x-y|}{\delta_n} \right)^{2-d} + O\left(\frac{|x-y|}{\delta_n} \right)^{-d} \right)$$

$$= a_d |x-y|^{2-d} + O(\delta_n^2 |x-y|^{-d}).$$

Likewise, in dimension two, using the standard estimate for the potential kernel (2.5) in (3.31) gives

$$g_n(x,y) = -\frac{2}{\pi} \log \frac{|x-y|}{\delta_n} + \frac{2}{\pi} \log \delta_n + O\left(\frac{|x-y|}{\delta_n} \right)^{-2}. \qquad \square$$

Our next result adapts Lemma 3.1.17(i) to the discrete setting. We list here our standing assumptions on the starting densities. Let σ be a function on $\mathbb{R}^d$ with compact support, such that

$$0 \le \sigma \le M \tag{3.32}$$

for some absolute constant M. Suppose that σ satisfies

$$\mathcal{L}\left(DC(\sigma) \right) = 0 \tag{3.33}$$

where $DC(\sigma)$ denotes the set of points in $\mathbb{R}^d$ where σ is discontinuous.

For $n = 1, 2, \ldots$ let σ_n be a function on $\delta_n \mathbb{Z}^d$ satisfying

$$0 \le \sigma_n \le M \tag{3.34}$$

and

$$\sigma_n^\square(x) \to \sigma(x), \qquad x \notin DC(\sigma). \tag{3.35}$$

Finally, suppose that

There is a ball $B \subset \mathbb{R}^d$ containing the supports of σ and $\sigma_n^\square$ for all n. $\qquad$ (3.36)

Although for maximum generality we have chosen to state our hypotheses on σ and σ_n separately, we remark that the above hypotheses on σ_n are satisfied in the particular case when σ_n is given by averaging σ over a small box:

$$\sigma_n(x) = \delta_n^{-d} \int_{x^\square} \sigma(y) dy. \tag{3.37}$$

In parallel to (3.3), for $x \in \delta_n \mathbb{Z}^d$ write

$$G_n \sigma_n(x) = \delta_n^d \sum_{y \in \delta_n \mathbb{Z}^d} g_n(x, y) \sigma_n(y), \tag{3.38}$$

where g_n is given by (3.30) and (3.31).

Lemma 3.1.23. *If σ, σ_n satisfy (3.32)-(3.36), then*

$$(G_n \sigma_n)^\square \to G\sigma$$

uniformly on compact subsets of $\mathbb{R}^d$.

Proof. Let $K \subset \mathbb{R}^d$ be compact. By the triangle inequality

$$|(G_n \sigma_n)^\square - G\sigma| \le |(G_n \sigma_n)^\square - G\sigma_n^\square| + |G\sigma_n^\square - G\sigma|.$$

By Lemma 3.1.17(i) the second term on the right side is $< \epsilon/2$ on K for sufficiently large n. The first term is at most

$$
\begin{aligned}
|(G_n \sigma_n)^\square(x) - G\sigma_n^\square(x)| &\le \sum_{y \in \delta_n \mathbb{Z}^d} \sigma_n(y) \left| \delta_n^d g_n(x^{\cdot\cdot}, y) - \int_{y^\square} g(x, z) dz \right| \\
&\le M \sum_{y \in B^{\cdot\cdot}} \int_{y^\square} |g_n(x^{\cdot\cdot}, y) - g(x, z)| dz \\
&= M \int_{B^{\cdot\cdot\square}} |g_n(x^{\cdot\cdot}, z^{\cdot\cdot}) - g(x, z)| dz. \tag{3.39}
\end{aligned}
$$

By Lemma 3.1.22, we have

$$|g_n(x^{\cdot\cdot}, z^{\cdot\cdot}) - g(x, z)| \leq |g_n(x^{\cdot\cdot}, z^{\cdot\cdot}) - g(x^{\cdot\cdot}, z^{\cdot\cdot})| + |g(x^{\cdot\cdot}, z^{\cdot\cdot}) - g(x, z)|$$
$$\leq C\delta_n|x - z|^{1-d}$$

for a constant C depending only on d. Integrating (3.39) in spherical shells about x, we obtain

$$|(G_n\sigma_n)^\square - G\sigma_n^\square| \leq Cd\omega_d MR\delta_n,$$

where R is the radius of B. Taking n large enough so that the right side is $< \epsilon/2$, the proof is complete. $\qquad\square$

Our next result adapts Lemma 3.1.9 to the discrete setting. Let σ_n be a function on $\delta_n\mathbb{Z}^d$ with finite support, and let γ_n be the function on $\delta_n\mathbb{Z}^d$ defined by

$$\gamma_n(x) = -|x|^2 - G_n\sigma_n(x). \tag{3.40}$$

Let

$$s_n = \inf\{f(x)|f \text{ is superharmonic on } \delta_n\mathbb{Z}^d \text{ and } f \geq \gamma_n\} \tag{3.41}$$

be the discrete least superharmonic majorant of γ_n, and let

$$D_n = \{x \in \delta_n \mathbb{Z}^d \,|s_n(x) > \gamma_n(x)\}. \tag{3.42}$$

Lemma 3.1.24. *Let γ_n, s_n, D_n be given by (3.40)-(3.42). If $\Omega \subset \delta_n\mathbb{Z}^d$ satisfies $D_n \subset \Omega$, then*

$$s_n(x) = \inf\{f(x)|f \text{ is superharmonic on } \Omega \text{ and } f \geq \gamma_n\}. \tag{3.43}$$

Proof. Let f be any function which is superharmonic on Ω and $\geq \gamma_n$. Since s_n is harmonic on D_n, so $f - s_n$ is superharmonic on D_n and attains its minimum in $D_n \cup \partial D_n$ on the boundary. Hence $f - s_n$ attains its minimum in $\Omega \cup \partial\Omega$ at a point x where $s_n(x) = \gamma_n(x)$. Since $f \geq \gamma_n$ we conclude that $f \geq s_n$ on Ω and hence everywhere. Thus s_n is at most the infimum in (3.43). Since the infimum in (3.43) is taken over a strictly larger set than that in (3.41), the reverse inequality is trivial. $\qquad\square$

3.2   Divisible Sandpile

3.2.1   Convergence of Odometers

By the odometer function for the divisible sandpile on $\delta_n \mathbb{Z}^d$ with source density σ_n, we will mean the function

$$u_n(x) = \delta_n^2 \cdot \text{total mass emitted from } x \text{ if each site } y \text{ starts with mass } \sigma_n(y).$$

Theorem 3.2.1. *Let u_n be the odometer function for the divisible sandpile on $\delta_n \mathbb{Z}^d$ with source density σ_n. If σ, σ_n satisfy (3.32)-(3.36), then*

$$u_n^\square \to s - \gamma \qquad \text{uniformly,}$$

where

$$\gamma(x) = -|x|^2 - G\sigma(x),$$

$G\sigma$ is given by (3.3), and s is the least superharmonic majorant of γ.

Lemma 3.2.2. *Let D_n be the set of fully occupied sites for the divisible sandpile in $\delta_n \mathbb{Z}^d$ started from source density σ_n. There is a ball $\Omega \subset \mathbb{R}^d$ with*

$$\bigcup_{n \geq 1} D_n^\square \cup D \subset \Omega.$$

Proof. Let A_n be the set of fully occupied sites for the divisible sandpile in $\delta_n \mathbb{Z}^d$ started from source density $\tau(x) = M 1_{x \in B}$, where B is given by (3.36). From the abelian property, Lemma 2.2.1, we have $D_n \subset A_n$. By the inner bound of Theorem 1.2.2 if we start with mass $m = 2\delta_n^{-d}\mathcal{L}(B)$ at the origin in $\delta_n \mathbb{Z}^d$, the resulting set of fully occupied sites contains $B^{\ddot{\ }}$; by the abelian property it follows that if we start with mass Mm at the origin in $\delta_n \mathbb{Z}^d$, the resulting set Ω_n of fully occupied sites contains A_n. By the outer bound of Theorem 1.2.2, Ω_n is contained in a ball Ω of volume $3M\mathcal{L}(B)$. By Lemma 3.1.16 we can enlarge Ω if necessary to contain D. $\qquad\qquad\square$

For $x \in \delta_n \mathbb{Z}^d$ write

$$\gamma_n(x) = -|x|^2 - G_n \sigma_n(x),$$

where G_n is defined by (3.38). Denote by s_n the least superharmonic majorant of γ_n in the lattice $\delta_n \mathbb{Z}^d$.

Lemma 3.2.3. *Let Ω be as in Lemma 3.2.2. There is a constant M' independent of n, such that $|\gamma_n| \leq M'$ in $\Omega^{\therefore}$.*

Proof. By Lemma 3.1.22 we have for $x \in \Omega^{\therefore}$

$$|G_n \sigma_n(x)| \leq \delta_n^d \sum_{y \in B^{\therefore}} |g_n(x,y)| \sigma_n(y)$$

$$\leq 2M \delta_n^d \sum_{y \in B^{\therefore}} |g(x,y)|$$

$$\leq CMR^2 \log R$$

for a constant C depending only on d, where R is the radius of Ω. It follows that $|\gamma_n| \leq (CM+1)R^2 \log R$ in $\Omega^{\therefore}$. $\qquad\square$

Lemma 3.2.4. *Fix $x \in \mathbb{R}^d$, and for $y \in \delta_n \mathbb{Z}^d$ let*

$$\alpha(y) = \delta_n^d g_n(x^{\therefore}, y) - \int_{y^{\square}} g(x,z)dz. \tag{3.44}$$

There are constants C_1, C_2 depending only on d, such that

  (i) $|\alpha(y)| \leq C_1 \delta_n^{1+d} |x-y|^{1-d}$.

  (ii) *If $y_1 \sim y_2$, then $|\alpha(y_1) - \alpha(y_2)| \leq C_2 \delta_n^{2+d} |x - y_1|^{-d}$.*

Proof. (i) By Lemma 3.1.22, if $z \in y^{\square}$ then

$$|g_n(x^{\therefore}, y) - g(x,z)| \leq |g_n(x^{\therefore}, y) - g(x^{\therefore}, y)| + |g(x^{\therefore}, y) - g(x, z)|$$

$$\leq C\delta_n |x - y|^{1-d}$$

for a constant C depending only on d. Integrating over $y^{\square}$ gives the result.

(ii) Let $p = y_2 - y_1$. By Lemma 3.1.22, we have

$$|\alpha(y_1) - \alpha(y_2)| \leq \int_{y_1^{\square}} |g_n(x^{\therefore}, y_1) - g_n(x^{\therefore}, y_2) - g(x, z) + g(x, z+p)| \, dz$$

$$\leq \int_{y_1^{\square}} |g(x^{\therefore}, y_1) - g(x^{\therefore}, y_2) - g(x, z) + g(x, z+p)| \, dz +$$

$$+ O(\delta_n^{d+2} |x - y_1|^{-d}).$$

Writing $q = x^{\therefore} - x - y_1 + z$, the quantity inside the integral can be expressed as

$$|f(o) - f(p) - f(q) + f(p+q)|,$$

where

$$f(w) = g(w, x^{\cdot\cdot} - y_1).$$

Since $|p|, |q|, |p+q| \le \delta_n(\sqrt{d}+1)$, we have by Taylor's theorem with remainder

$$|f(o) - f(p) - f(q) + f(p+q)| \le 3d\delta_n^2 \sum_{i,j=1}^{d} \left| \frac{\partial^2 f}{\partial x_i \partial x_j}(o) \right|$$

$$\le \begin{cases} 6d^2(d-1)(d-2)a_d\delta_n^2 |x-y_1|^{-d}, & d \ge 3 \\ (24/\pi)\delta_n^2 |x-y_1|^{-2}, & d = 2. \quad \square \end{cases}$$

Lemma 3.2.5. *Let Ω be an open ball as in Lemma 3.2.2, and let Ω_1 be a ball with $\bar{\Omega} \subset \Omega_1$. Let*

$$\phi_n = -\Delta(s_n 1_{\Omega_{\ddot{1}}}).$$

Then

$$\left| s_n^{\square} - G\phi_n^{\square} \right| \to 0$$

uniformly on Ω.

Proof. Let $\nu_n(x)$ be the amount of mass present at x in the final state of the divisible sandpile on $\delta_n \mathbb{Z}^d$. By Lemma 2.2.2 we have $s_n = u_n + \gamma_n$, hence

$$\Delta s_n = (\nu_n - \sigma_n) + (\sigma_n - 1) = \nu_n - 1.$$

In particular, $|\Delta s_n| \le 1$.

Since

$$G\phi_n^{\square}(x) = \sum_{y \in \Omega_{\ddot{1}} \cup \partial\Omega_{\ddot{1}}} \phi_n(y) \int_{y^{\square}} g(x, z) dz$$

and $s_n = G_n\phi_n$ in $\Omega^{\cdot\cdot}$, we have for $x \in \Omega$

$$G\phi_n^{\square}(x) - s_n^{\square}(x) = G\phi_n^{\square}(x) - G_n\phi_n(x^{\cdot\cdot})$$

$$= -\sum_{y \in \Omega_{\ddot{1}} \cup \partial\Omega_{\ddot{1}}} \phi_n(y)\alpha(y)$$

where $\alpha(y)$ is given by (3.44). Hence

$$G\phi_n^{\square}(x) - s_n^{\square}(x) = \sum_{y \in \Omega_{\ddot{1}}} \Delta s_n(y)\alpha(y) + \delta_n^{-2} \sum_{\substack{y \in \Omega_{\ddot{1}}, z \notin \Omega_{\ddot{1}} \\ y \sim z}} (s_n(y)\alpha(z) - s_n(z)\alpha(y)). \qquad (3.45)$$

Let R, R_1 be the radii of Ω, Ω_1. By Lemma 3.2.4(i), summing in spherical shells about x, the first sum in (3.45) is bounded in absolute value by

$$\sum_{y \in \Omega_{\ddot{1}}} |\alpha(y)| \leq \int_0^{2R_1} \frac{d\omega_d r^{d-1}}{\delta_n^d} (C_1 \delta_n^{1+d} r^{1-d}) dr = C_1 d\omega_d R_1 \delta_n.$$

To bound the second sum in (3.45), note that $s_n = \gamma_n$ outside Ω, so

$$|s_n(y)\alpha(z) - s_n(z)\alpha(y)| \leq |\gamma_n(y)||\alpha(y) - \alpha(z)| + |\alpha(y)||\gamma_n(y) - \gamma_n(z)|. \tag{3.46}$$

By Lemmas 3.2.3 and 3.2.4(ii), the first term is bounded by

$$|\gamma_n(y)||\alpha(y) - \alpha(z)| \leq C_2 M' \delta_n^{2+d} |x - y|^{-d}.$$

Fix $\epsilon > 0$, and let Ω_2 be a ball with $\bar{\Omega}_1 \subset \Omega_2$. Since γ is uniformly continuous on Ω_2, and $\gamma_n \to \gamma$ uniformly on Ω_2 by Lemma 3.1.23, for sufficiently large n we have

$$|\gamma_n(y) - \gamma_n(z)| \leq |\gamma_n(y) - \gamma(y)| + |\gamma(y) - \gamma(z)| + |\gamma(z) - \gamma_n(z)| \leq \epsilon.$$

Thus by Lemma 3.2.4(i) the second term in (3.46) is bounded by

$$|\alpha(y)||\gamma_n(y) - \gamma_n(z)| \leq C_1 \delta_n^{1+d} |x - y|^{1-d} \epsilon.$$

Since $x \in \Omega$ and y is adjacent to $\partial \Omega_1$, we have $|x - y| \geq R_1 - R - \delta_n$, so the second term in (3.45) is bounded in absolute value by

$$2\delta_n^{-2} \# \partial \Omega_{\ddot{1}} \left(C_2 M' \delta_n^{2+d} (R_1 - R)^{-d} + C_1 \delta_n^{1+d} (R_1 - R)^{1-d} \epsilon \right) \leq C_3 \epsilon$$

for sufficiently large n. $\qquad\square$

Lemma 3.2.6. $s_n^{\square} \to s$ *uniformly on compact subsets of* $\mathbb{R}^d$.

Proof. By Lemma 3.2.2 there is a ball Ω containing D and $D_n^{\square}$ for all n. Outside Ω we have

$$s_n^{\square} = \gamma_n^{\square} \to \gamma = s$$

uniformly on compact sets by Lemma 3.1.23.

To show convergence in Ω, write

$$\tilde{s}(x) = \int_{\mathbb{R}^d} s(y) \lambda^{-d} \eta\left(\frac{x - y}{\lambda}\right) dy, \tag{3.47}$$

where η is the standard smooth mollifier

$$\eta(x) = \begin{cases} Ce^{1/(|x|^2-1)}, & |x| < 1 \\ 0, & |x| \geq 1 \end{cases}$$

normalized so that $\int_{\mathbb{R}^d} \eta \, dx = 1$ (see [17, Appendix C.4]). Then $\tilde{s}$ is smooth and superharmonic. Fix $\epsilon > 0$. By Lemma 3.1.2(ii) and compactness, s is uniformly continuous on $\bar{\Omega}$, so taking λ sufficiently small in (3.47) we have $|s - \tilde{s}| < \epsilon$ in Ω. Let A_ϵ be the maximum third partial of $\tilde{s}$ in Ω. By Lemma 3.1.21 the function

$$q_n(x) = \tilde{s}^{\cdots}(x) - \frac{1}{6} A_\epsilon \delta_n |x|^2$$

is superharmonic in $\Omega^{\cdots}$. By Lemma 3.1.23 we have $\gamma_n^\square \to \gamma$ uniformly in Ω. Taking N large enough so that $\frac{1}{6} A_\epsilon \delta_n |x|^2 < \epsilon$ in Ω and $|\gamma_n - \gamma^{\cdots}| < \epsilon$ in $\Omega^{\cdots}$ for all $n > N$, we obtain

$$q_n > \tilde{s}^{\cdots} - \epsilon > s^{\cdots} - 2\epsilon \geq \gamma^{\cdots} - 2\epsilon > \gamma_n - 3\epsilon$$

in $\Omega^{\cdots}$. In particular, the function $f_n = \max(q_n + 3\epsilon, \gamma_n)$ is superharmonic in $\Omega^{\cdots}$. By Lemma 3.1.24 it follows that $f_n \geq s_n$, hence

$$s_n \leq q_n + 3\epsilon < \tilde{s}^{\cdots} + 3\epsilon < s^{\cdots} + 4\epsilon$$

in $\Omega^{\cdots}$. By the uniform continuity of s on $\bar{\Omega}$, taking N larger if necessary we have $|s - s^{\cdots\square}| < \epsilon$ in Ω, and hence $s_n^\square < s + 5\epsilon$ in Ω for all $n > N$.

For the reverse inequality, let

$$\phi_n = -\Delta(s_n \mathbf{1}_{\Omega_1^{\cdots}})$$

where Ω_1 is an open ball containing $\bar{\Omega}$. By Lemma 3.2.5 we have

$$\left| s_n^\square - G\phi_n^\square \right| < \epsilon$$

and hence

$$G\phi_n^\square > \gamma_n^\square - \epsilon > \gamma - 2\epsilon. \tag{3.48}$$

on Ω for sufficiently large n. Since $\phi_n^\square$ is nonnegative on Ω, by Lemma 3.1.4 the function $G\phi_n^\square$ is superharmonic on Ω, so by (3.48) the function

$$\psi_n = \max(G\phi_n^\square + 2\epsilon, \gamma)$$

is superharmonic on Ω for sufficiently large n. By Lemma 3.1.9 it follows that $\psi_n \geq s$, hence

$$s_n^{\square} > G\phi_n^{\square} - \epsilon \geq s - 3\epsilon$$

on Ω for sufficiently large n. $\square$

Proof of Theorem 3.2.1. Let Ω be as in Lemma 3.2.2. By Lemmas 3.1.23 and 3.2.6 we have $\gamma_n^{\square} \to \gamma$ and $s_n^{\square} \to s$ uniformly on Ω. By Lemma 2.2.2, we have $u_n = s_n - \gamma_n$. Since $u_n^{\square} = 0 = s - \gamma$ off Ω, we conclude that $u_n^{\square} \to s - \gamma$ uniformly. $\square$

3.2.2 Convergence of Domains

In addition to the conditions (3.32)-(3.36) assumed in the previous section, we assume in this section that the source density σ satisfies

$$\text{For all } x \in \mathbb{R}^d \text{ either } \sigma(x) \leq \lambda \text{ or } \sigma(x) \geq 1 \tag{3.49}$$

for a constant $\lambda < 1$. We also assume that

$$\{\sigma \geq 1\} = \overline{\{\sigma \geq 1\}^o}. \tag{3.50}$$

Moreover, we assume that for any $\epsilon > 0$ there exists $N(\epsilon)$ such that

$$\text{If } x \in \{\sigma \geq 1\}_\epsilon, \text{ then } \sigma_n(x) \geq 1 \text{ for all } n \geq N(\epsilon); \tag{3.51}$$

and

$$\text{If } x \notin \{\sigma \geq 1\}^\epsilon, \text{ then } \sigma_n(x) \leq \lambda \text{ for all } n \geq N(\epsilon). \tag{3.52}$$

As before, we have chosen to state the hypotheses on σ and σ_n separately for maximum generality, but all hypotheses on σ_n are satisfied in the particular case when σ_n is given by averaging σ in a small box (3.37).

We set

$$\gamma(x) = -|x|^2 - G\sigma(x)$$

with s the least superharmonic majorant of γ and

$$D = \{x \in \mathbb{R}^d | s(x) > \gamma(x)\}.$$

We also write

$$\widetilde{D} = D \cup \{x \in \mathbb{R}^d | \sigma(x) \geq 1\}^o.$$

For a domain $A \subset \mathbb{R}^d$, denote by A_ϵ and A^ϵ its inner and outer open ϵ-neighborhoods, respectively.

Theorem 3.2.7. *Let σ and σ_n satisfy (3.32)-(3.36) and (3.49)-(3.52). For $n \geq 1$ let D_n be the domain of fully occupied sites for the divisible sandpile in $\delta_n \mathbb{Z}^d$ started from source density σ_n. For any $\epsilon > 0$ we have for large enough n*

$$\widetilde{D}^{::}_\epsilon \subset D_n \subset \widetilde{D}^{\epsilon::}. \tag{3.53}$$

According to the following lemma, near any occupied site $x \in D_n$ lying outside $\widetilde{D}^\epsilon$, we can find a site y where the odometer u_n is relatively large. This is a discrete version of a standard argument for the obstacle problem; see for example Friedman [19], Ch. 2 Lemma 3.1.

Lemma 3.2.8. *Fix $\epsilon > 0$ and $x \in D_n$ with $x \notin \widetilde{D}^\epsilon$. If n is sufficiently large, there is a point $y \in \delta_n \mathbb{Z}^d$ with $|x - y| \leq \frac{\epsilon}{2} + \delta_n$ and*

$$u_n(y) \geq u_n(x) + \frac{1 - \lambda}{4} \epsilon^2.$$

Proof. By (3.50), we have $\{\sigma \geq 1\} \subset \overline{\widetilde{D}}$. Thus if $x \notin \widetilde{D}^\epsilon$, the ball $B = B(x, \frac{\epsilon}{2})^{::}$ is disjoint from $\{\sigma \geq 1\}^{\epsilon/2}$. In particular, if $n \geq N(\frac{\epsilon}{2})$, then by (3.52) we have $\sigma_n \leq \lambda$ on B. Thus the function

$$w(y) = u_n(y) - (1 - \lambda)|x - y|^2$$

is subharmonic on $B \cap D_n$, so it attains its maximum on the boundary. Since $w(x) \geq 0$, the maximum cannot be attained on ∂D_n, where u_n vanishes; so it is attained at some point $y \in \partial B$, and

$$u_n(y) \geq w(y) + (1 - \lambda) \left(\frac{\epsilon}{2}\right)^2.$$

Since $w(y) \geq w(x) = u_n(x)$, the proof is complete. $\qquad \square$

Proof of Theorem 3.2.7. Fix $\epsilon > 0$. By Lemma 3.1.20 we have

$$\widetilde{D}_\epsilon \subset D_\eta \cup \{\sigma \geq 1\}_\eta$$

for some $\eta > 0$. Let $u = s - \gamma$. Since the closure of D_η is compact and contained in D, we have $u \geq m_\eta$ on D_η for some $m_\eta > 0$. By Theorem 3.2.1, for sufficiently large n we have $u_n > u - \frac{1}{2} m_\eta > 0$ on $D_\eta^{::}$, hence $D_\eta^{::} \subset D_n$. Likewise, by (3.51) we have $\{\sigma \geq 1\}_\eta^{::} \subset D_n$ for large enough n. Thus $\widetilde{D}_\epsilon \subset D_n$.

For the other inclusion, fix $x \in \delta_n \mathbb{Z}^d$ with $x \notin \widetilde{D}^\epsilon$. Since u vanishes on the ball $B = B(x, \frac{\epsilon}{2})$, by Theorem 3.2.1 we have $u_n < \frac{1-\lambda}{4} \epsilon^2$ on $B^{::} \cup \partial B^{::}$ for all sufficiently large n. By Lemma 3.2.8 we conclude that $x \notin D_n$, and hence $D_n \subset \widetilde{D}^{\epsilon::}$. $\qquad \square$

3.3 Rotor-Router Model

In trying to adapt the proofs of Theorems 3.2.1 and 3.2.7 to the rotor-router model, we are faced with two main problems. The first is to define an appropriate notion of convergence of integer-valued densities σ_n on $\delta_n \mathbb{Z}^d$ to a real-valued density σ on $\mathbb{R}^d$. The requirement that σ_n take only integer values is of course imposed on us by the nature of the rotor-router model itself, since unlike the divisible sandpile, the rotor-router works with discrete, indivisible particles. The second problem is to find an appropriate analogue of Lemma 2.2.2 for the rotor-router model.

Although these two problems may seem unrelated, the single technique of *smoothing* neatly takes care of them both. To illustrate the basic idea, suppose we are given a domain $A \subset \mathbb{R}^d$, and let σ_n be the function on $\delta_n \mathbb{Z}^d$ taking the value 1 on odd lattice points in $A^{\cdot\cdot}$, and the value 2 on even lattice points in $A^{\cdot\cdot}$, while vanishing outside $A^{\cdot\cdot}$. We would like to find a sense in which σ_n converges to the real-valued density $\sigma = \frac{3}{2} 1_A$. One approach is to average σ_n in a box whose side length L_n goes to zero more slowly than the lattice spacing: $L_n \downarrow 0$ while $L_n/\delta_n \uparrow \infty$ as $n \uparrow \infty$. The resulting "smoothed" version of σ_n converges to σ pointwise away from the boundary of A.

By smoothing the odometer function in the same way, we can obtain an approximate analogue of Lemma 2.2.2 for the rotor-router model. Rather than average in a box as described above, however, it is technically more convenient to average according to the distribution of a lazy random walk run for a fixed number $\alpha(n)$ of steps. Denote by $(X_k)_{k \geq 0}$ the lazy random walk in $\delta_n \mathbb{Z}^d$ which stays in place with probability $\frac{1}{2}$ and moves to each of the $2d$ neighbors with probability $\frac{1}{4d}$. Given a function f on $\delta_n \mathbb{Z}^d$, define its *k-smoothing*

$$S_k f(x) = \mathbb{E}(X_k | X_0 = x). \tag{3.54}$$

From the Markov property we have $S_k S_\ell = S_{k+\ell}$. Also, the discrete Laplacian can be written as

$$\Delta = 2\delta_n^{-2}(S_1 - S_0). \tag{3.55}$$

In particular, $\Delta S_k = S_k \Delta$.

3.3.1 Convergence of Odometers

For $n = 1, 2, \ldots$ let σ_n be an integer-valued function on $\delta_n \mathbb{Z}^d$ satisfying $0 \leq \sigma_n \leq M$. We assume as usual that there is a ball $B \subset \mathbb{R}^d$ containing the support of σ_n for all n.

Let σ be a function on $\mathbb{R}^d$ supported in B satisfying (3.32) and (3.33). In place of condition (3.35) we assume that there exist integers $\alpha(n) \uparrow \infty$ with $\delta_n \alpha(n) \downarrow 0$ such that

$$(S_{\alpha(n)}\sigma_n)^\square(x) \to \sigma(x), \qquad x \notin DC(\sigma). \tag{3.56}$$

By the odometer function for rotor-router aggregation starting from source density σ_n, we will mean the function

$$u_n(x) = \delta_n^2 \cdot \text{number of particles emitted from } x$$

if $\sigma_n(y)$ particles start at each site y.

Theorem 3.3.1. *Let u_n be the odometer function for rotor-router aggregation on $\delta_n \mathbb{Z}^d$ starting from source density σ_n. If σ, σ_n satisfy (3.32)-(3.34), (3.36) and (3.56), then $u_n^\square \to s - \gamma$ uniformly, where*

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

and s is the least superharmonic majorant of γ.

Given a function f on $\delta_n \mathbb{Z}^d$, for an edge (x,y) write

$$\nabla f(x,y) = \frac{f(y) - f(x)}{\delta_n}.$$

Given a function κ on edges in $\delta_n \mathbb{Z}^d$, write

$$\text{div } \kappa(x) = \frac{1}{2d\delta_n} \sum_{y \sim x} \kappa(x,y).$$

The discrete Laplacian on $\delta_n \mathbb{Z}^d$ is then given by

$$\Delta f(x) = \text{div } \nabla f = \delta_n^{-2}\left(\frac{1}{2d}\sum_{y \sim x} f(y) - f(x)\right).$$

The following "rescaled" version of Lemma 2.4.1 is proved in the same way.

Lemma 3.3.2. *For an edge (x,y) in $\delta_n \mathbb{Z}^d$, denote by $\kappa(x,y)$ the net number of crossings from x to y performed by particles during a sequence of rotor-router moves. Let*

$$u(x) = \delta_n^2 \cdot \text{number of particles emitted from } x \text{ during this sequence.}$$

Then

$$\nabla u(x,y) = \delta_n(-2d\kappa(x,y) + \rho(x,y)). \tag{3.57}$$

for some edge function ρ which satisfies

$$|\rho(x,y)| \le 4d - 2$$

independent of x, y, n and the chosen sequence of moves.

Proof. Writing $N(x,y)$ for the number of particles routed from x to y, for any $y, z \sim x$ we have

$$|N(x,y) - N(x,z)| \le 1.$$

Since $u(x) = \delta_n^2 \sum_{y \sim x} N(x,y)$, we obtain

$$\delta_n^{-2} u(x) - 2d + 1 \le 2dN(x,y) \le \delta_n^{-2} u(x) + 2d - 1$$

hence

$$|\nabla u(x,y) + 2d\delta_n \kappa(x,y)| = \delta_n |\delta_n^{-2} u(y) - \delta_n^{-2} u(x) + 2dN(x,y) - 2dN(y,x)|$$
$$\le (4d - 2)\delta_n. \quad \square$$

Recall that the divisible sandpile odometer function has Laplacian $1 - \sigma_n$ inside the set D_n of fully occupied sites. The next lemma shows that the same is approximately true of the smoothed rotor-router odometer function. Denote by R_n the occupied shape for rotor-router aggregation on $\delta_n \mathbb{Z}^d$ starting from source density σ_n.

Lemma 3.3.3. $|\Delta S_k u_n(x) - \mathbb{P}_x(X_k \in R_n) + S_k \sigma_n(x)| \le C_0/\sqrt{k+1}.$

Proof. Let κ and ρ be defined as in Lemma 3.3.2. Since each site x starts with $\sigma_n(x)$ particles and ends with either one particle or none accordingly as $x \in R_n$ or $x \notin R_n$, we have

$$2d\delta_n \operatorname{div} \kappa = \sigma_n - 1_{R_n}.$$

Taking the divergence in (3.57) we obtain

$$\Delta u_n = 1_{R_n} - \sigma_n + \delta_n \operatorname{div} \rho. \tag{3.58}$$

Using the fact that S_k and Δ commute, it follows that

$$\Delta S_k u_n(x) = \mathbb{P}_x(X_k \in R_n) - \mathbb{E}_x \sigma_n(X_k) + \delta_n \mathbb{E}_x \operatorname{div} \rho(X_k). \tag{3.59}$$

Since ∇u and κ are antisymmetric, ρ is antisymmetric by (3.57). Thus the last term in (3.59) can be written

$$\delta_n \, \mathbb{E}_x \, \text{div} \, \rho(X_k) = \frac{1}{2d} \sum_{y \sim z} \mathbb{P}_x(X_k = y) R(y, z)$$

$$= -\frac{1}{2d} \sum_{y \sim z} \mathbb{P}_x(X_k = z) \rho(y, z)$$

$$= \frac{1}{4d} \sum_{y \sim z} (\mathbb{P}_x(X_k = y) - \mathbb{P}_x(X_k = z)) \rho(y, z) \qquad (3.60)$$

where the sums are taken over all pairs of neighboring sites $y, z \in \delta_n \mathbb{Z}^d$.

We can couple lazy random walk X_k started at x with lazy random walk X_k' started at a uniform neighbor of x so that the probability of not coupling in k steps is at most $C/\sqrt{k+1}$, where C is a constant depending only on d [36]. Since the total variation distance between X_k and X_k' is at most the probability of not coupling, we obtain

$$\sum_{y \sim z} | \, \mathbb{P}_x \, (X_k = y) - \mathbb{P}_x(X_k = z)| \leq \frac{C}{\sqrt{k+1}}.$$

Using the fact that $\rho(y, z)$ is uniformly bounded in (3.60), taking $C_0 = 4Cd$ completes the proof. $\qquad \square$

The next lemma shows that smoothing the odometer function does not introduce much extra error.

Lemma 3.3.4. $|S_k u_n - u_n| \leq \delta_n^2(\frac{1}{2} M k + C_0 \sqrt{k+1})$.

Proof. From (3.55) we have

$$|S_k u_n - u_n| \leq \sum_{j=0}^{k-1} |S_{j+1} u_n - S_j u_n|$$

$$= \frac{\delta_n^2}{2} \sum_{j=0}^{k-1} |\Delta S_j u_n|.$$

But by Lemma 3.3.3

$$|\Delta S_j u_n| \leq M + \frac{C_0}{\sqrt{j+1}}.$$

Summing over j yields the result. $\qquad \square$

Let

$$\gamma_n(x) = -|x|^2 - G_n S_{\alpha(n)} \sigma_n(x)$$

and let s_n be the least superharmonic majorant of γ_n. By Lemma 2.2.2, the difference $s_n - \gamma_n$ is the odometer function for the divisible sandpile on $\delta_n \mathbb{Z}^d$ starting from the smoothed source density $\widetilde{\sigma}_n = S_{\alpha(n)} \sigma_n$. Note that by Lemma 3.2.2, there is a ball $\Omega \subset \mathbb{R}^d$ containing the supports of $s - \gamma$ and of $s_n - \gamma_n$ for all n. The next lemma compares the smoothed rotor-router odometer for the source density σ_n with the divisible sandpile odometer for the smoothed density $\widetilde{\sigma}_n$.

Lemma 3.3.5. *Let $\Omega \subset \mathbb{R}^d$ be a ball centered at the origin containing the supports of $s - \gamma$ and of $s_n - \gamma_n$ for all n. Then*

$$S_{\alpha(n)} u_n \geq s_n - \gamma_n - C_0 r^2 \alpha(n)^{-1/2} \tag{3.61}$$

on all of $\delta_n \mathbb{Z}^d$, where r is the radius of Ω.

Proof. Since

$$\Delta \gamma_n = -1 + S_{\alpha(n)} \sigma_n,$$

by Lemma 3.3.3 the function

$$f(x) = S_{\alpha(n)} u_n(x) + \gamma_n(x) + C_0 \alpha(n)^{-1/2}(r^2 - |x|^2)$$

is superharmonic on $\delta_n \mathbb{Z}^d$. Since $f \geq \gamma_n$ on $\Omega^{\because}$, the function $\psi_n = \max(f, \gamma_n)$ is superharmonic on $\Omega^{\because}$, hence $\psi_n \geq s_n$ by Lemma 3.1.24. Thus $f \geq s_n$ on $\Omega^{\because}$, so (3.61) holds on $\Omega^{\because}$ and hence everywhere. $\qquad \square$

Lemma 3.3.6. *Let R_n be the set of occupied sites for rotor-router aggregation in $\delta_n \mathbb{Z}^d$ started from source density σ_n. There is a ball $\Omega \subset \mathbb{R}^d$ with $\bigcup R_n^{\square} \subset \Omega$.*

Proof. By assumption there is a ball $B \subset \mathbb{R}^d$ containing the support of σ_n for all n. Let A_n be the set of occupied sites for rotor-router aggregation in $\delta_n \mathbb{Z}^d$ started from source density $\tau(x) = M 1_{x \in B}$. From the abelian property we have $R_n \subset A_n$. By the inner bound of Theorem 1.2.1, if we start with $\lfloor 2 \delta_n^{-d} \mathcal{L}(B) \rfloor$ particles at the origin in $\delta_n \mathbb{Z}^d$, the resulting set of occupied sites contains $B^{\because}$; by the abelian property it follows that if we start with $M \lfloor 2 \delta_n^{-d} \mathcal{L}(B) \rfloor$ particles at the origin, the resulting set Ω_n of fully occupied sites contains A_n. By the outer bound of Theorem 1.2.1, Ω_n is contained in a ball Ω of volume $3 M \mathcal{L}(B)$. $\qquad \square$

Proof of Theorem 3.3.1. By Lemma 3.2.2 there is a ball Ω containing the support of $s - \gamma$ and of $s_n - \gamma_n$ for all n. By Lemma 3.3.6 we can enlarge Ω if necessary to contain the support of $S_{\alpha(n)} u_n$ for all n. By Lemma 3.3.3, the function

$$\phi(x) = S_{\alpha(n)} u_n(x) - s_n(x) + \gamma_n(x) + C_0 \alpha(n)^{-1/2} |x|^2.$$

is subharmonic on the set

$$\widetilde{R}_n = \{x \in R_n : y \in R_n \text{ whenever } ||x - y||_1 \leq \alpha(n)\}$$

since $\mathbb{P}_x \left(X_{\alpha(n)} \in R_n \right) = 1$ for $x \in \widetilde{R}_n$. From Lemma 3.3.3 we have $\left| \Delta S_{\alpha(n)} u_n \right| \leq M + C_0$, so for $x \notin \widetilde{R}_n$ we have by Lemma 3.1.18

$$S_{\alpha(n)} u_n(x) \leq c(M + C_0) \alpha(n)^2 \delta_n^2.$$

By the maximum principle in $\widetilde{R}_n$

$$\phi(x) \leq c(M + C_0) \alpha(n)^2 \delta_n^2 + C_0 \alpha(n)^{-1/2} r^2, \qquad x \in \widetilde{R}_n.$$

From Lemma 3.3.5 we obtain

$$- C_0 r^2 \alpha(n)^{-1/2} \leq S_{\alpha(n)} u_n - s_n + \gamma_n \leq c(M + C_0) \alpha(n)^2 \delta_n^2 + C_0 r^2 \alpha(n)^{-1/2} \qquad (3.62)$$

on all of $\delta_n \mathbb{Z}^d$, where r is the radius of Ω.

By Lemmas 3.1.23 and 3.2.6 we have $\gamma_n^{\square} \to \gamma$ and $s_n^{\square} \to s$ uniformly on Ω. Since $\alpha(n) \uparrow \infty$ and $\delta_n \alpha(n) \downarrow 0$, we conclude from (3.62) that $\left(S_{\alpha(n)} u_n \right)^{\square} \to s - \gamma$ uniformly on Ω. Since both $S_{\alpha(n)} u_n$ and $s - \gamma$ vanish outside Ω, this convergence is uniform on $\mathbb{R}^d$. By Lemma 3.3.4 we have

$$\left| S_{\alpha(n)} u_n - u_n \right| \to 0$$

uniformly on $\delta_n \mathbb{Z}^d$, and hence $u_n^{\square} \to s - \gamma$ uniformly on $\mathbb{R}^d$. $\qquad \square$

3.3.2 Convergence of Domains

In addition to the assumptions of the previous section, in this section we require that

$$\text{For all } x \in \mathbb{R}^d \text{ either } \sigma(x) \geq 1 \text{ or } \sigma(x) = 0. \qquad (3.63)$$

We also assume that

$$\{\sigma \geq 1\} = \overline{\{\sigma \geq 1\}^o}. \qquad (3.64)$$

Moreover, we assume that for any $\epsilon > 0$ there exists $N(\epsilon)$ such that

$$\text{If } x \in \{\sigma \geq 1\}_\epsilon, \text{ then } \sigma_n(x) \geq 1 \text{ for all } n \geq N(\epsilon); \tag{3.65}$$

and

$$\text{If } x \notin \{\sigma \geq 1\}^\epsilon, \text{ then } \sigma_n(x) = 0 \text{ for all } n \geq N(\epsilon). \tag{3.66}$$

Theorem 3.3.7. *Let σ and σ_n satisfy (3.32)-(3.34), (3.36), (3.56) and (3.63)-(3.66). For $n \geq 1$ let $R_n = \{u_n > 0\}$ be the domain of sites in $\delta_n \mathbb{Z}^d$ that emit a particle in the rotor-router model started from source density σ_n. For any $\epsilon > 0$ we have for all sufficiently large n*

$$\widetilde{D}_\epsilon^{\cdot\cdot} \subset R_n \subset \widetilde{D}^{\epsilon\cdot\cdot}, \tag{3.67}$$

where

$$\widetilde{D} = \{s > \gamma\} \cup \{\sigma \geq 1\}^o.$$

Lemma 3.3.8. *Fix $\epsilon > 0$ and $n \geq N(\epsilon/2)$. Given $x \notin \widetilde{D}^\epsilon$ and $\delta_n \leq \rho \leq \epsilon/2$, let*

$$\mathcal{N}_\rho(x) = \#B(x, \rho) \cap R_n.$$

If $u_n \leq \delta_n^2 m$ on $B(x, \rho)$, then

$$\mathcal{N}_\rho(x) \geq \frac{m}{m-1} \mathcal{N}_{\rho - \delta_n}(x).$$

Proof. Since at least $\mathcal{N}_\rho(x)$ particles must enter the ball $B(x, \rho)$, we have

$$\sum_{y \in \partial B(x, \rho - \delta_n)} u_n(y) \geq \delta_n^2 \mathcal{N}_\rho(x).$$

There are $\mathcal{N}_\rho(x) - \mathcal{N}_{\rho - \delta_n}(x)$ terms in the sum on the left side, and each term is at most $\delta_n^2 m$, hence

$$m\big(\mathcal{N}_\rho(x) - \mathcal{N}_{\rho - \delta_n}(x)\big) \geq \mathcal{N}_\rho(x). \quad \square$$

The following lemma can be seen as a weak analogue of Lemma 3.2.8 for the divisible sandpile. In Lemma 3.3.11, below, we obtain a more exact analogue under slightly stronger hypotheses.

Lemma 3.3.9. *For any $\epsilon > 0$, if n is sufficiently large and $x \in R_n$ with $x \notin \widetilde{D}^\epsilon$, then there is a point $y \in R_n$ with $|x - y| \le \epsilon/2$ and*

$$u_n(y) \ge \frac{\epsilon \delta_n}{4 \log \left(2 \omega_d (\epsilon / 2 \delta_n)^d \right)}.$$

Proof. Take $n \ge N(\epsilon/2)$ large enough so that $4 \delta_n < \epsilon$. Let m be the maximum value of $\delta_n^{-2} u_n$ on $B(x, \epsilon)$. Since $x \in R_n$ we have $\mathcal{N}_0(x) = 1$. Iteratively applying Lemma 3.3.8, we obtain

$$\mathcal{N}_{\epsilon/2}(x) \ge \left(\frac{m}{m-1} \right)^{\lfloor \epsilon/2\delta_n \rfloor} \mathcal{N}_0(x) \ge \exp\left(\frac{\epsilon}{4m\delta_n} \right).$$

Since

$$\mathcal{N}_{\epsilon/2}(x) \le \#B(x, \epsilon/2) \le 2\omega_d \left(\frac{\epsilon}{2\delta_n} \right)^d,$$

we conclude that

$$\frac{\epsilon}{4m\delta_n} \le \log \left(2\omega_d (\epsilon/\delta_n)^d \right).$$

Solving for m yields the result. $\qquad\square$

The following lemma shows that far away from D, the rotor-router odometer grows at most quadratically as we move away from the boundary of R_n.

Lemma 3.3.10. *For any $\epsilon > 0$ and $k \ge 1$, if n is sufficiently large and $z \notin R_n \cup \widetilde{D}^\epsilon$, then*

$$u_n(x) \le 4cdk^2\delta_n^2, \qquad x \in B(z, k\delta_n)^{\because},$$

where $c = c_{1/2}$ is the constant in Lemma 3.1.18.

Proof. Taking the divergence in (3.57), since $-2d\delta_n \operatorname{div} \kappa(x)$ is the net number of particles entering x, we obtain

$$\Delta u_n = 1_{R_n} - \sigma_n + \delta_n \operatorname{div} R. \tag{3.68}$$

By (3.66), if $n \ge N(\epsilon/2)$ is large enough so that $k\delta_n < \epsilon/4$, we have $\sigma_n = 0$ on the ball $B = B(z, 2k\delta_n)^{\because}$. Since $|R| \le 4d - 2$, we have $|\Delta u_n| \le 4d$ on B, so by Lemma 3.1.18

$$u_n(x) \le 4cd|x - z|^2$$

for $x \in B(z, k\delta_n)^{\because}$. $\qquad\square$

The following lemma is analogous to Lemma 3.2.8 for the divisible sandpile.

Lemma 3.3.11. *Fix $\epsilon > 0$ and $k \geq 4C_0^2$, where C_0 is the constant in Lemma 3.3.3. There is a constant C_1 such that if $x \in R_n$, $x \notin \widetilde{D}^\epsilon$ satisfies*

$$S_k u_n(x) > C_1 k^2 \delta_n^2$$

and n is sufficiently large, there exists $y \in \delta_n \mathbb{Z}^d$ with $|x - y| \leq \frac{\epsilon}{2} + \delta_n$ and

$$S_k u_n(y) \geq S_k u_n(x) + \frac{\epsilon^2}{8}.$$

Proof. Let

$$R_n^{(k)} = \{x \in R_n : y \in R_n \text{ whenever } \|x - y\|_1 \leq k\delta_n\}.$$

By Lemma 3.3.3, we have

$$\Delta S_k u_n(y) \geq 1 - S_k \sigma_n(y) - C_0/\sqrt{k}, \qquad y \in R_n^{(k)}.$$

Note that $C_0/\sqrt{k} < \frac{1}{2}$. Take $n \geq N(\epsilon/4)$ large enough so that $k\delta_n < \epsilon/4$; then $S_k \sigma_n$ vanishes on the ball $B = B(x, \epsilon/2)^{\cdot\cdot}$ by (3.66). Thus the function

$$f(y) = S_k u_n(y) - \frac{1}{2}|x - y|^2$$

is subharmonic on $\widetilde{R}_n := R_n^{(k)} \cap B$, so it attains its maximum on the boundary.

If $y \in \partial R_n^{(k)}$, there is a point $z \notin R_n$ with $|y - z| \leq k\delta_n$. By Lemma 3.3.3 we have $|\Delta u_n| \leq M + C_0$, hence by Lemma 3.1.18 it follows that

$$u_n \leq 4c(M + C_0)k^2\delta_n^2$$

in the ball $B(z, 2k\delta_n)$. Taking $C_1 = 4c(M + C_0)$ we find that

$$S_k u_n(y) \leq C_1 k^2 \delta_n^2 < S_k u_n(x).$$

Thus f cannot attain its maximum in $\widetilde{R}_n$ on $\partial R_n^{(k)}$, and hence must attain its maximum at a point $y \in \partial B$. Since $f(y) \geq f(x)$ we conclude that

$$\begin{aligned}
S_k u_n(y) &= f(y) + \frac{1}{2}|x - y|^2 \\
&\geq S_k u_n(x) + \frac{1}{2}\left(\frac{\epsilon}{2}\right)^2. \qquad \square
\end{aligned}$$

Proof of Theorem 3.3.7. Fix $\epsilon > 0$. By Lemma 3.1.20 we have

$$\widetilde{D}_\epsilon \subset D_\eta \cup \{\sigma \geq 1\}_\eta$$

for some $\eta > 0$. Since the closure of D_η is compact and contained in D, we have $u \geq m_\eta$ on D_η for some $m_\eta > 0$. By Theorem 3.3.1, for sufficiently large n we have $u_n > u - \frac{1}{2}m_\eta > 0$ on $D_\eta^{\because}$, hence $D_\eta^{\because} \subset R_n$. Likewise, by (3.51) we have $\{\sigma \geq 1\}_\eta^{\because} \subset R_n$ for large enough n. Thus $\widetilde{D}_\epsilon^{\because} \subset R_n$.

For the other inclusion, fix $x \in \delta_n \mathbb{Z}^d$ with $x \notin \widetilde{D}^\epsilon$. Since u vanishes on the ball $B = B(x, \epsilon)$, by Theorem 3.3.1 we have $u_n < \frac{\epsilon^2}{8}$ on $B^{\because}$ for all sufficiently large n. Let $k = 4C_0^2$, and take n large enough so that $k\delta_n < \epsilon/4$. Then $S_k u_n < \frac{\epsilon^2}{8}$ on the ball $B(x, 3\epsilon/4)^{\because}$. By Lemma 3.3.11 it follows that $S_k u_n \leq C_1 k^2 \delta_n^2$ on the smaller ball $B' = B(x, \epsilon/4)^{\because}$. In particular, for $y \in B'$ we have

$$u_n(y) \leq \frac{S_k u_n(y)}{\mathbb{P}_y(X_k = y)} < C_2 k^{d/2+2} \delta_n^2.$$

For sufficiently large n the right side is at most $\epsilon \delta_n / 4 \log \left(2\omega_d (\epsilon/2\delta_n)^d \right)$, so we conclude from Lemma 3.3.9 that $x \notin R_n$. $\qquad\square$

3.4   Internal DLA

Our hypotheses for convergence of internal DLA domains are the same as those for the rotor-router model: σ is a bounded, nonnegative, compactly supported function on $\mathbb{R}^d$ that is continuous almost everywhere, and $\{\sigma_n\}_{n \geq 1}$ is a sequence of uniformly bounded functions on $\delta_n \mathbb{Z}^d$ with uniformly bounded supports, whose "smoothings" converge pointwise to σ at all continuity points of σ.

Theorem 3.4.1. *Let σ and σ_n satisfy (3.32)-(3.34), (3.36), (3.56) and (3.63)-(3.66). For $n \geq 1$ let I_n be the random domain of occupied sites for internal DLA in $\delta_n \mathbb{Z}^d$ started from source density σ_n. For all $\epsilon > 0$ we have with probability one*

$$\widetilde{D}_\epsilon^{\because} \subset I_n \subset \widetilde{D}^{\epsilon\because} \qquad \text{for all sufficiently large } n, \tag{3.69}$$

where

$$\widetilde{D} = \{s > \gamma\} \cup \{\sigma \geq 1\}^o.$$

3.4.1 Inner Estimate

Fix $n \geq 1$, and label the particles in $\delta_n \mathbb{Z}^d$ by the integers $1, \ldots, m_n$, where $m_n = \sum_{x \in \delta_n \mathbb{Z}^d} \sigma_n(x)$. Let x_i be the starting location of the particle labeled i, so that

$$\#\{i \,|\, x_i = x\} = \sigma_n(x).$$

For each $i = 1, \ldots, m_n$ let $(X_t^i)_{t \geq 0}$ be a simple random walk in $\delta_n \mathbb{Z}^d$ such that $X_0^i = x_i$, with X^i and X^j independent for $i \neq j$.

For $z \in \delta_n \mathbb{Z}^d$ and $\epsilon > 0$, consider the stopping times

$$\tau_z^i = \inf\left\{t \geq 0 \,|\, X_t^i = z\right\};$$
$$\tau_\epsilon^i = \inf\left\{t \geq 0 \,|\, X_t^i \notin D_\epsilon^{\cdots}\right\};$$
$$\nu^i = \inf\left\{t \geq 0 \,|\, X_t^i \notin \{X_{\nu^j}^j\}_{j=1}^{i-1}\right\}. \tag{3.70}$$

The stopping time ν^i is defined inductively in i with $\nu^1 = 0$. We think of building up the internal DLA cluster one site at a time, by letting the particle labeled i walk until it exits the set of sites already occupied by particles with smaller labels. Thus ν^i is the number of steps taken by the particle labeled i, and $X_{\nu^i}^i$ is the location where it stops.

Fix $z \in D_\epsilon^{\cdots}$ and consider the random variables

$$\mathcal{M}_\epsilon = \sum_{i=1}^{m_n} 1_{\{\tau_z^i < \tau_\epsilon^i\}};$$
$$L_\epsilon = \sum_{i=1}^{m_n} 1_{\{\nu^i \leq \tau_z^i < \tau_\epsilon^i\}}.$$

These sums can be interpreted in terms of the following two-stage procedure. During the first stage, we form the occupied internal DLA cluster by adding on one site at a time, as described above. In the second stage, we let each particle continue walking from where it stopped until it exits $D_\epsilon^{\cdots}$. Then $\mathcal{M}_\epsilon$ counts the total number of particles that visit z during both stages, while L_ϵ counts the number of particles that visit z during the second stage. In particular, if $L_\epsilon < \mathcal{M}_\epsilon$, then z was visited during the first stage and hence belongs to the occupied cluster.

The sum L_ϵ is difficult to estimate directly because the indicator random variables in the sum are not independent. Following [29], we can bound L_ϵ by a sum of independent indicators as follows. For each site $y \in D_\epsilon^{\cdots}$, let $(Y_t^y)_{t \geq 0}$ be a simple random walk in $\delta_n \mathbb{Z}^d$

such that $Y_0^y = y$, with Y^x and Y^y independent for $x \neq y$. Let

$$\widetilde{L}_\epsilon = \sum_{y \in D_\epsilon^{\vdots}} 1_{\{\tau_z^y < \tau_\epsilon^y\}}$$

where

$$\tau_z^y = \inf\left\{t \geq 0 \,|\, Y_t^y = z\right\};$$
$$\tau_\epsilon^y = \inf\left\{t \geq 0 \,|\, Y_t^y \notin D_\epsilon^{\vdots}\right\}.$$

Thus $\widetilde{L}_\epsilon$ counts the number of walks Y^y that hit z before exiting $D_\epsilon^{\vdots}$. Since the sites $X_{\nu^i}^i$ are distinct, we can couple the walks $\{Y^y\}$ and $\{X^i\}$ so that $L_\epsilon \leq \widetilde{L}_\epsilon$.

Define

$$f_{n,\epsilon}(z) = g_{n,\epsilon}(z,z) \, \mathbb{E}\left(\mathcal{M}_\epsilon - \widetilde{L}_\epsilon\right),$$

where

$$g_{n,\epsilon}(y,z) = \mathbb{E}\#\{t < \tau_\epsilon^y \,|\, Y_t^y = z\}$$

is the Green's function for simple random walk in $\delta_n \mathbb{Z}^d$ stopped on exiting $D_\epsilon^{\vdots}$. Then

$$f_{n,\epsilon}(z) = g_{n,\epsilon}(z,z) \left(\sum_{i=1}^{m_n} \mathbb{P}\left(\tau_z^i < \tau_\epsilon^i\right) - \sum_{y \in D_\epsilon^{\vdots}} \mathbb{P}\left(\tau_z^y < \tau_\epsilon^y\right)\right)$$

$$= g_{n,\epsilon}(z,z) \sum_{y \in D_\epsilon^{\vdots}} \left(\sigma_n(y) - 1\right) \mathbb{P}\left(\tau_z^y < \tau_\epsilon^y\right)$$

$$= \sum_{y \in D_\epsilon^{\vdots}} \left(\sigma_n(y) - 1\right) g_{n,\epsilon}(y,z) \tag{3.71}$$

where in the last step we have used the identity

$$\mathbb{P}\left(\tau_z^y < \tau_\epsilon^y\right) = \frac{g_{n,\epsilon}(y,z)}{g_{n,\epsilon}(z,z)}. \tag{3.72}$$

Thus $f_{n,\epsilon}$ solves the Dirichlet problem

$$\Delta f_{n,\epsilon} = \delta_n^{-2}(1 - \sigma_n), \qquad\qquad \text{on } D_\epsilon^{\vdots}; \tag{3.73}$$

$$f_{n,\epsilon} = 0, \qquad\qquad \text{on } \partial D_\epsilon^{\vdots}.$$

Note that the divisible sandpile odometer function u_n for the source density σ_n solves exactly the same Dirichlet problem, with the domain $D_\epsilon^{\vdots}$ in (3.73) replaced by the domain D_n of fully occupied sites. Our strategy will be first to use Theorem 3.2.7 to argue

that since $D_n \to D$, the solutions to the Dirichlet problems in D_n and $D_\epsilon^{\cdot\cdot}$ should be close; next, by Theorem 3.2.1, since $u_n \to u$ it follows that the functions $f_{n,\epsilon}$ and u are close. Since u is strictly positive in D_ϵ, we obtain in this way a lower bound on $f_{n,\epsilon}$, and hence a lower bound on $\mathbb{E}(\mathcal{M}_\epsilon - \widetilde{L}_\epsilon)$. Finally, using large deviations for sums of independent indicators, we conclude that with high probability $\widetilde{L}_\epsilon < \mathcal{M}_\epsilon$, and hence with high probability every point $z \in D_\epsilon^{\cdot\cdot}$ belongs to the occupied cluster I_n. The core of the argument is Lemma 3.4.3, below, which gives the desired lower bound on $f_{n,\epsilon}$.

In order to apply Theorems 3.2.1 and 3.2.7, we must have discrete densities σ_n which converge pointwise to σ at all continuity points of σ. Recall, however, that in order to allow σ to assume non-integer values, we have chosen not to assume that σ_n converges to σ, but rather only that the *smoothed density* $S_{\alpha(n)}\sigma_n$ converges to σ; see (3.56). For this reason, we will need to run the divisible sandpile on $\delta_n \mathbb{Z}^d$ starting from the smoothed density rather than from σ_n, and we will use the following smoothed version of equation (3.73).

$$\Delta S_k f_{n,\epsilon} = \delta_n^{-2}(1 - S_k \sigma_n), \qquad \text{on } D_{\epsilon'}^{\cdot\cdot}. \tag{3.74}$$

Here $k \geq 1$ is arbitrary, and $\epsilon' = \epsilon + k\delta_n$.

The following lemma is proved in the same way as Lemma 3.3.4.

Lemma 3.4.2. *Let $\epsilon' = \epsilon + k\delta_n$. Then for $z \in D_{\epsilon'}^{\cdot\cdot}$ we have*

$$|S_k f_{n,\epsilon}(z) - f_{n,\epsilon}(z)| \leq \frac{1}{2}Mk.$$

Proof. From (3.55) and (3.74) we have for $z \in D_{\epsilon'}^{\cdot\cdot}$

$$
\begin{aligned}
|S_k f_{n,\epsilon}(z) - f_{n,\epsilon}(z)| &\leq \sum_{j=0}^{k-1} |S_{j+1} f_{n,\epsilon}(z) - S_j f_{n,\epsilon}(z)| \\
&= \frac{\delta_n^2}{2} \sum_{j=0}^{k-1} |\Delta S_j f_{n,\epsilon}(z)| \\
&\leq \frac{1}{2}Mk. \quad \square
\end{aligned}
$$

Lemma 3.4.3. *Fix $\epsilon > 0$, and let $\beta > 0$ be the minimum value of $u = s - \gamma$ on $\overline{D_\epsilon}$. There exists $0 < \eta < \epsilon$ such that for all sufficiently large n*

$$f_{n,\eta}(z) \geq \frac{1}{2}\beta\delta_n^{-2}, \qquad z \in D_\epsilon^{\cdot\cdot}.$$

Proof. Since u is uniformly continuous on $\overline{D}$, we can choose $\eta > 0$ small enough so that $u \leq \beta/8$ outside $D_{2\eta}$. Let $\eta' = \eta + \delta_n \alpha(n)$. Since $\delta_n \alpha(n) \downarrow 0$, we have $u \leq \beta/8$ outside $D_{\eta'}$ if n is sufficiently large. Since $u \geq \beta$ on $\overline{D_\epsilon}$, for large enough n we have $\partial D_{\eta'} \not\subset \overline{D_\epsilon}$ and hence $\eta' < \epsilon$. Let u_n be the odometer function for the divisible sandpile on $\delta_n \mathbb{Z}^d$ started from source density $S_{\alpha(n)} \sigma_n$. We have $u_n^{\square} \to u$ uniformly by Theorem 3.2.1, so for n sufficiently large we have $|u_n - u^{\cdot\cdot}| \leq \beta/8$, hence $u_n \leq \beta/4$ on $\partial D_{\eta'}^{\cdot\cdot}$.

Let D_n be the domain of fully occupied sites for the divisible sandpile in $\delta_n \mathbb{Z}^d$ started from source density $S_{\alpha(n)} \sigma_n$. By Theorem 3.2.7 we have $D_\eta^{\cdot\cdot} \subset D_n$ for all sufficiently large n. Now (3.73) implies that $S_{\alpha(n)} f_{n,\eta} - \delta_n^{-2} u_n$ is harmonic on $D_{\eta'}^{\cdot\cdot}$. By Lemmas 3.4.2 and 3.1.18, for $z \in \partial D_{\eta'}^{\cdot\cdot}$ we have

$$|S_{\alpha(n)} f_{n,\eta}(z)| \leq |f_{n,\eta}(z)| + \frac{1}{2} M\alpha(n)$$
$$\leq cM\alpha(n)^2 + \frac{1}{2} M\alpha(n).$$

Since $\delta_n \alpha(n) \downarrow 0$, for sufficiently large n we have

$$S_{\alpha(n)} f_{n,\eta}(z) \geq -\frac{\beta}{16\delta_n^2}, \qquad z \in \partial D_{\eta'}^{\cdot\cdot}.$$

Since $S_{\alpha(n)} f_{n,\eta} - \delta_n^{-2} u_n$ attains its minimum in $D_{\eta'}^{\cdot\cdot}$ on the boundary, for $z \in D_{\eta'}^{\cdot\cdot}$ we have

$$S_{\alpha(n)} f_{n,\eta}(z) \geq \delta_n^{-2} \left(u_n(z) - \frac{5\beta}{16} \right)$$
$$\geq \delta_n^{-2} \left(u(z) - \frac{7\beta}{16} \right).$$

Since $u \geq \beta$ in $D_\epsilon^{\cdot\cdot}$, taking n large enough so that $\alpha(n)\delta_n^2 \leq \beta/8M$, we conclude from Lemma 3.4.2 that

$$f_{n,\eta} \geq \frac{9\beta}{16\delta_n^2} - \frac{1}{2} M\alpha(n) \geq \frac{1}{2}\beta\delta_n^{-2}.$$

on $D_\epsilon^{\cdot\cdot}$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ $\square$

Lemma 3.4.4. *We have*

$$\mathbb{E}\widetilde{L}_\eta = \frac{\mathbb{E}\tau_\eta^z}{g_{n,\eta}(z,z)}.$$

Moreover

$$\mathbb{E}\mathcal{M}_\eta \leq M \frac{\mathbb{E}\tau_\eta^z}{g_{n,\eta}(z,z)}.$$

Proof. Let

$$T_z^y = \inf\{t \geq 0 | Y_t^y = z\}$$

be the first hitting time of z for the walk Y^y. By (3.72) and the symmetry of $g_{n,\eta}$, we have

$$\mathbb{E}\widetilde{L}_\eta = \sum_{y \in D_{\tilde{\eta}}^{\ddot{.}}} \mathbb{P}\big(T_z^y < \tau_\eta^y\big) \tag{3.75}$$

$$= \sum_{y \in D_{\tilde{\eta}}^{\ddot{.}}} \frac{g_{n,\eta}(y, z)}{g_{n,\eta}(z, z)}$$

$$= \frac{1}{g_{n,\eta}(z, z)} \sum_{y \in D_{\tilde{\eta}}^{\ddot{.}}} g_{n,\eta}(z, y).$$

The sum in the last line is $\mathbb{E}_z T_{\partial D_{\tilde{\eta}}^{\ddot{.}}}$.

To prove the inequality for $\mathbb{E}\mathcal{M}_\eta$, note that $\mathbb{E}\mathcal{M}_\eta$ is bounded above by M times the sum on the right side of (3.75). $\qquad\square$

The next lemma, using a martingale argument to compute the expected time for simple random walk to exit a ball, is well known.

Lemma 3.4.5. *Fix $r > 0$ and let $B = B(o, r)^{\ddot{:}} \subset \delta_n \mathbb{Z}^d$, and let T be the first hitting time of ∂B. Then*

$$\mathbb{E}_o T = \left(\frac{r}{\delta_n}\right)^2 + O\left(\frac{r}{\delta_n}\right).$$

Proof. Since $\delta_n^{-2}|X_t|^2 - t$ is a martingale with bounded increments, and $\mathbb{E}_o T < \infty$, by optional stopping we have

$$\mathbb{E}_o T = \delta_n^{-2} \, \mathbb{E}_o \, |X_T|^2 = \delta_n^{-2}(r + O(\delta_n))^2. \quad \square$$

The next lemma, which bounds the expected number of times simple random walk returns to the origin before reaching distance r, is also well known.

Lemma 3.4.6. *Fix $r > 0$ and let $B = B(o, r)^{\ddot{:}} \subset \delta_n \mathbb{Z}^d$, and let G_B be the Green's function for simple random walk stopped on exiting B. If n is sufficiently large, then*

$$G_B(o, o) \leq \log \frac{r}{\delta_n}.$$

Proof. In dimension $d \geq 3$ the result is trivial since simple random walk on $\delta_n \mathbb{Z}^d$ is transient. In dimension two, consider the function

$$f(x) = G_B(o, x) - g_n(o, x)$$

where g_n is the rescaled potential kernel defined in (3.31). Since f is harmonic in B it attains its maximum on the boundary, hence by Lemma 3.1.22 we have for $x \in B$

$$f(x) \leq \frac{2}{\pi} \log r + O\left(\frac{\delta_n^2}{r^2}\right).$$

Since $g_n(o, o) = \frac{2}{\pi} \log \delta_n$, the result follows on setting $x = o$. $\qquad\square$

We will use the following large deviation bound; for a proof, see [2, Cor. A.14].

Lemma 3.4.7. *If N is a sum of finitely many independent indicator random variables, then for all $\lambda > 0$*

$$\mathbb{P}(|N - \mathbb{E}N| > \lambda \, \mathbb{E} \, N) < 2e^{-c_\lambda \mathbb{E}N}$$

where $c_\lambda > 0$ is a constant depending only on λ.

Let

$$\tilde{I}_n = \left\{ X^i_{\nu^i} | \nu^i < \tilde{\tau}^i \right\} \subset I_n$$

where ν^i is given by (3.70), and

$$\tilde{\tau}^i = \inf \left\{ t \geq 0 | X^i_t \notin \widetilde{D}^{::} \right\}.$$

The inner estimate of Theorem 3.4.1 follows immediately from the lemma below. Although for the inner estimate it suffices to prove Lemma 3.4.8 with I_n in place of $\tilde{I}_n$, we will make use of the stronger statement with $\tilde{I}_n$ in the proof of the outer estimate in the next section.

Lemma 3.4.8. *For any $\epsilon > 0$,*

$$\mathbb{P}\left(\widetilde{D}^{::}_\epsilon \subset \tilde{I}_n \text{ for all but finitely many } n\right) = 1.$$

Proof. For $z \in \widetilde{D}^{::}_\epsilon$, let $\mathcal{E}_z(n)$ be the event that $z \notin \tilde{I}_n$. By Borel-Cantelli it suffices to show that

$$\sum_{n \geq 1} \sum_{z \in \widetilde{D}^{::}_\epsilon} \mathbb{P}(\mathcal{E}_z(n)) < \infty. \tag{3.76}$$

By Lemma 3.1.20, since $\widetilde{D} = D \cup \{\sigma \geq 1\}^o$ we have

$$\widetilde{D}_\epsilon \subset D_{\epsilon'} \cup \{\sigma \geq 1\}_{\epsilon'}$$

for some $\epsilon' > 0$. By (3.65), for $n \geq N(\epsilon')$ the terms in (3.76) with $z \in \{\sigma \geq 1\}_{\epsilon'}$ vanish, so it suffices to show

$$\sum_{n \geq 1} \sum_{z \in D_{\epsilon'}^{\cdots}} \mathbb{P}(\mathcal{E}_z(n)) < \infty. \tag{3.77}$$

By Lemma 3.4.3 there exists $0 < \eta < \epsilon'$ such that

$$f_{n,\eta}(z) \geq \frac{1}{2}\beta\delta_n^{-2}, \qquad z \in D_{\epsilon'}^{\cdots} \tag{3.78}$$

for all sufficiently large n, where $\beta > 0$ is the minimum value of u on $\overline{D_{\epsilon'}}$. Fixing $z \in D_{\epsilon'}^{\cdots}$, since $L_\eta \leq \widetilde{L}_\eta$ we have

$$\begin{aligned} \mathbb{P}(\mathcal{E}_z(n)) &\leq \mathbb{P}(\mathcal{M}_\eta = L_\eta) \\ &\leq \mathbb{P}(\mathcal{M}_\eta \leq \widetilde{L}_\eta) \\ &\leq \mathbb{P}(\mathcal{M}_\eta \leq a) + \mathbb{P}(\widetilde{L}_\eta \geq a) \end{aligned} \tag{3.79}$$

for a real number a to be chosen below. By Lemma 3.4.7, since $\widetilde{L}_\eta$ and $\mathcal{M}_\eta$ are sums of independent indicators, we have

$$\mathbb{P}\left(\widetilde{L}_\eta \geq (1+\lambda)\,\mathbb{E}\,\widetilde{L}_\eta\right) < 2e^{-c_\lambda \mathbb{E}\widetilde{L}_\eta} \tag{3.80}$$

$$\mathbb{P}\left(\mathcal{M}_\eta \leq (1-\lambda)\,\mathbb{E}\,\mathcal{M}_\eta\right) < 2e^{-c_\lambda \mathbb{E}\mathcal{M}_\eta}$$

where c_λ depends only on λ, chosen below. Now since $z \in D_{\epsilon'}$ and D is bounded by Lemma 3.1.16, we have

$$B(z, \epsilon' - \eta) \subset D_\eta \subset B(z, R)$$

hence by Lemma 3.4.5

$$\frac{1}{2}\left(\frac{\epsilon' - \eta}{\delta_n}\right)^2 \leq \mathbb{E}\tau_\eta^z \leq \left(\frac{R}{\delta_n}\right)^2. \tag{3.81}$$

By Lemma 3.4.4 it follows that $\mathbb{E}\mathcal{M}_\eta \leq MR^2/\delta_n^2 g_{n,\eta}(z,z)$. Taking

$$a = \mathbb{E}\widetilde{L}_\eta + \frac{\beta}{4\delta_n^2 g_{n,\eta}(z,z)}$$

in (3.79), and letting $\lambda = \beta/4MR^2$ in (3.80), we have

$$\lambda\,\mathbb{E}\,\widetilde{L}_\eta \leq \lambda\,\mathbb{E}\,\mathcal{M}_\eta \leq \frac{\beta}{4\delta_n^2 g_{n,\eta}(z,z)}, \tag{3.82}$$

hence

$$a \geq (1 + \lambda) \, \mathbb{E} \, \widetilde{L}_\eta.$$

Moreover, from (3.78) we have

$$\mathbb{E}\mathcal{M}_\eta - E\widetilde{L}_\eta = \frac{f_{n,\eta}(z)}{g_{n,\eta}(z,z)} \geq \frac{\beta}{2\delta_n^2 g_{n,\eta}(z,z)},$$

hence by (3.82)

$$a \leq \mathbb{E}\mathcal{M}_\eta - \frac{\beta}{4\delta_n^2 g_{n,\eta}(z,z)} \leq (1 - \lambda) \, \mathbb{E} \, \mathcal{M}_\eta.$$

Thus we obtain from (3.79) and (3.80)

$$\mathbb{P} \left(\mathcal{E}_z(n) \right) \leq 4e^{-c_\lambda \mathbb{E}\widetilde{L}_\eta}. \tag{3.83}$$

By Lemmas 3.4.4 and 3.4.6 along with (3.81)

$$\mathbb{E}\widetilde{L}_\eta = \frac{\mathbb{E}_z T_{\partial D_{\widetilde{\eta}}^{\cdots}}}{g_{n,\eta}(z,z)}$$

$$\geq \frac{1}{2} \left(\frac{\epsilon' - \eta}{\delta_n} \right)^2 \frac{1}{\log(R/\delta_n)}.$$

Using (3.83), the sum in (3.77) is thus bounded by

$$\sum_{n \geq 1} \sum_{z \in D_{\epsilon'}^{\cdots}} \mathbb{P}(\mathcal{E}_z(n)) \leq \sum_{n \geq 1} \delta_n^{-d} \omega_d R^d \cdot 4 \exp \left(-\frac{c_\lambda (\epsilon' - \eta)^2}{2\delta_n^2 \log(R/\delta_n)} \right) < \infty. \quad \square$$

3.4.2 Outer Estimate

For $x \in \mathbb{Z}^d$ write

$$Q(x, h) = \{y \in \mathbb{Z}^d : ||x - y||_\infty \leq h\}$$

for the cube of side length $2h + 1$ centered at x. According to the next lemma, if we start a simple random walk at distance h from a hyperplane $H \subset \mathbb{Z}^d$, then the walk is farily "spread out" by the time it hits H, in the sense that its chance of first hitting H at any particular point z has the same order of magnitude as z ranges over a $(d - 1)$-dimensional cube of side length order h.

Lemma 3.4.9. *Fix $y \in \mathbb{Z}^d$ with $y_1 = h$, and let T be the first hitting time of the hyperplane $H = \{x \in \mathbb{Z}^d | x_1 = 0\}$. Let $F = H \cap Q(y, h)$. For any $z \in F$ we have*

$$\mathbb{P}_y(X_T = z) \geq ah^{1-d}$$

for a constant a depending only on d.

Proof. For fixed $w \in H$, the function

$$f(x) = \mathbb{P}_x(X_T = w)$$

is harmonic in the ball $B = B(y, h)$. By the Harnack inequality [28, Theorem 1.7.2] we have

$$f(x) \geq cf(y), \qquad x \in B(y, h/2)$$

for a constant c depending only on d. By translation invariance, it follows that for $w' \in H \cap B(w, h/2)$, letting $x = y + w - w'$ we have

$$\mathbb{P}_y(X_T = w') = f(x) \geq cf(y) = c\,\mathbb{P}_y(X_T = w).$$

Iterating, we obtain for any $w' \in H \cap Q(w, h)$

$$\mathbb{P}_y(X_T = w') \geq c^{\sqrt{d}}\,\mathbb{P}_y(X_T = w). \tag{3.84}$$

Let T' be the first exit time of the cube $Q(y, h-1)$. Since F is a boundary face of this cube, we have $\{X_{T'} \in F\} \subset \{X_T \in F\}$. Let z_0 be the closest point to y in H. Taking $w' = z_0$, we have $w' \in H \cap Q(w, h)$ whenever $w \in F$. Summing (3.84) over $w \in F$, we obtain

$$(2h+1)^{d-1}c^{-\sqrt{d}}\,\mathbb{P}_y(X_T = z_0) \geq \mathbb{P}_y(X_T \in F) \geq \mathbb{P}(X_{T'} \in F) = \frac{1}{2d}.$$

Now for any $z \in F$, taking $w = z_0$ and $w' = z$ in (3.84), we conclude that

$$\mathbb{P}_y(X_T = z) \geq c^{\sqrt{d}}\,\mathbb{P}_y(X_T = z_0) \geq \frac{c^{2\sqrt{d}}}{2d}(2h+1)^{1-d}. \qquad \square$$

Lemma 3.4.10. *Let h, ρ be positive integers. Let N be the number of particles that ever visit the cube $Q(o, \rho)$ during the internal DLA process, if one particle starts at each site $y \in \mathbb{Z}^d - Q(o, \rho + h)$, and k additional particles start at sites $y_1, \ldots, y_k \notin Q(o, \rho + h)$. If $k \leq \frac{1}{4}h^d$, then*

$$N \leq \mathrm{Binom}(k, p),$$

where $p < 1$ is a constant depending only on d.

Proof. Let F_j be a half-space defining a face of the cube $Q = Q(o, \rho)$, such that $\text{dist}(y_j, F_j) \geq h$. Let $z_j \in F_j$ be the closest point to y_j in F_j, and let $B_j = Q(z_j, h/2) \cap F_j^c$. Let Z_j be the random set of sites where the particles starting at $y_1, \ldots, y_j$ stop. Since $\#B_j \geq \frac{1}{2}h^d \geq 2k$, there is a hyperplane H_j parallel to F_j and intersecting B_j, such that

$$\#B_j \cap H_j \cap Z_{j-1} \leq \frac{1}{2}h^{d-1}. \tag{3.85}$$

Denote by A_j the event that the particle starting at y_j ever visits Q. On this event, the particle must pass through H_j at a site which was already occupied by an earlier particle. Thus if $\{X_t\}_{t \geq 0}$ is the random walk performed by the particle starting at y_j, then

$$A_j \subset \{X_T \in Z_{j-1}\},$$

where T is the first hitting time of H_j. By Lemma 3.4.9, every site $z \in B_j \cap H_j$ satisfies

$$\mathbb{P}(X_T = z) \geq ah^{1-d}.$$

From (3.85), since $\#B_j \cap H_j \geq h^{d-1}$ we obtain $\mathbb{P}(X_T \notin Z_{j-1}) \geq a/2$, hence

$$\mathbb{P}(A_j | \mathcal{F}_{j-1}) \leq p$$

where $p = 1 - a/2$, and $\mathcal{F}_i$ is the σ-algebra generated by the walks performed by the particles starting at $y_1, \ldots, y_i$. Thus we can couple the indicators 1_{A_j} with i.i.d. indicators $I_j \geq 1_{A_j}$ of mean p to obtain

$$N = \sum_{j=1}^{k} 1_{A_j} \leq \sum_{j=1}^{k} I_j = \text{Binom}(k, p). \quad \square$$

The next lemma shows that if few enough particles start outside a cube $Q(o, 3\rho)$, it is highly unlikely that any of them will reach the smaller cube $Q(o, \rho)$.

Lemma 3.4.11. *Let ρ, k be positive integers with*

$$k \leq \frac{1}{4}\left(1 - p^{1/2d}\right)^d \rho^d$$

where $p < 1$ is the constant in Lemma 3.4.10. Let N be the number of particles that ever visit the cube $Q(o, \rho)$ during the internal DLA process, if one particle starts at each site $y \in \mathbb{Z}^d - Q(o, 3\rho)$, and k additional particles start at sites $y_1, \ldots, y_k \notin Q(o, 3\rho)$. Then

$$\mathbb{P}(N > 0) \leq c_0 e^{-c_1 \rho}$$

where $c_0, c_1 > 0$ are constants depending only on d.

Proof. Let N_j be the number of particles that ever visit the cube $Q_j = Q(o, \rho_j)$, where

$$\rho_j = \left(2 + p^{j/2d}\right)\rho.$$

Let $k_j = p^{j/2}k$, and let A_j be the event that $N_j \leq k_j$. Taking $h = \rho_j - \rho_{j+1}$ in Lemma 3.4.10, since

$$k_j \leq \frac{1}{4}p^{j/2}\left(1 - p^{1/2d}\right)^d \rho^d = \frac{1}{4}h^d$$

we obtain

$$N_{j+1}1_{A_j} \leq \mathrm{Binom}(N_j, p).$$

Hence

$$\mathbb{P}(A_{j+1}|A_j) \geq \mathbb{P}\big(\mathrm{Binom}(k_j, p) \leq k_{j+1}\big)$$
$$\geq 1 - 2e^{-ck_j} \tag{3.86}$$

where in the second line we have used Lemma 3.4.7 with $\lambda = \sqrt{p} - p$.

Now let

$$j = \left\lfloor 2\frac{\log k - \log \rho}{\log(1/p)}\right\rfloor$$

so that $p^{1/2}\rho \leq k_j \leq \rho$. On the event A_j, at most ρ particles visit the cube $Q(o, 2\rho)$. Since the first particle to visit each cube $Q(o, 2\rho - i)$ stops there, at most $\rho - i$ particles visit $Q(o, 2\rho - i)$. Taking $i = \rho$ we obtain $\mathbb{P}(N = 0) \geq \mathbb{P}(A_j)$. From (3.86) we conclude that

$$\mathbb{P}(N = 0) \geq \mathbb{P}(A_1)\,\mathbb{P}\,(A_2|A_1)\cdots\mathbb{P}\,(A_j|A_{j-1}) \geq 1 - 2je^{-c\rho}.$$

The right side is at least $1 - c_0 e^{-c_1\rho}$ for suitable constants c_0, c_1. $\qquad\square$

Proof of Theorem 3.4.1. The inner estimate is immediate from Lemma 3.4.8. For the outer estimate, let

$$N_n = \#\big\{1 \leq i \leq m_n \big| \nu^i \geq \tilde{\tau}^i\big\} = m_n - \#\tilde{I}_n$$

be the number of particles that leave $\widetilde{D}^{\cdots}$ before aggregating to the cluster. Let $K_0 = \frac{1}{4}\left(\frac{2(1-p^{1/2d})}{3\sqrt{d}}\right)^d$, where p is the constant in Lemma 3.4.10. For $\epsilon > 0$ and $n_0 \geq 1$, consider the event

$$F_{n_0} = \big\{N_n \leq K_0(\epsilon/\delta_n)^d \text{ for all } n \geq n_0\big\}.$$

By (3.63) and (3.64), the closure of $\widetilde{D}$ contains the support of σ, so by Proposition 3.1.13,

$$\delta_n^d m_n = \delta_n^d \sum_{x \in \delta_n \mathbb{Z}^d} \sigma_n(x) \to \int_{\mathbb{R}^d} \sigma(x)dx = \mathcal{L}(\widetilde{D}). \tag{3.87}$$

Moreover, by Proposition 3.1.13(i), for sufficiently small η we have

$$\mathcal{L}\big(\widetilde{D} - \widetilde{D}_{2\eta}\big) \leq \frac{1}{2}K_0\epsilon^d.$$

Taking n large enough so that $\delta_n < \eta$, we obtain

$$\delta_n^d \# \widetilde{D}_\eta^{\vdots} = \mathcal{L}\big(\widetilde{D}_\eta^{\vdots\square}\big) \geq \mathcal{L}\big(\widetilde{D}_{2\eta}\big) \geq \mathcal{L}\big(\widetilde{D}\big) - \frac{1}{2}K_0\epsilon^d.$$

Thus for sufficiently large n, on the event $\widetilde{D}_\eta^{\vdots} \subset \tilde{I}_n$ we have

$$\begin{aligned} N_n &\leq m_n - \#\widetilde{D}_\eta^{\vdots} \\ &\leq m_n - \delta_n^{-d}\mathcal{L}\big(\widetilde{D}\big) + \frac{1}{2}K_0\epsilon^d\delta_n^{-d} \\ &\leq K_0\epsilon^d\delta_n^{-d}, \end{aligned}$$

where in the last line we have used (3.87). From Lemma 3.4.8 we obtain

$$\mathbb{P}\left(F_{n_0}\right) \geq \mathbb{P}\big(\widetilde{D}_\eta^{\vdots} \subseteq \tilde{I}_n \text{ for all } n \geq n_0\big) \uparrow 1 \tag{3.88}$$

as $n_0 \uparrow \infty$.

By compactness, we can find finitely many cubes $Q_1, \ldots, Q_m$ of side length $\rho = 2\epsilon/3\sqrt{d}$ centered at points in $\partial(\widetilde{D}^\epsilon)$, with $\partial(\widetilde{D}^\epsilon) \subset \bigcup Q_i$. Taking $k = \lfloor K_0(\epsilon/\delta_n)^d \rfloor$ in Lemma 3.4.11, since $3Q_i$ is disjoint from $\widetilde{D}$, we obtain for $n \geq n_0$

$$\mathbb{P}\big(\{Q_i \cap I_n \neq \emptyset\} \cap F_{n_0}\big) \leq c_0 e^{-c_1\rho}.$$

Summing over i yields

$$\mathbb{P}\big(\{\widetilde{D}^\epsilon \cap I_n \neq \emptyset\} \cap F_{n_0}\big) \leq c_0 m e^{-c_1\rho}.$$

By Borel-Cantelli, if G is the event that $I_n \not\subset \widetilde{D}^{\epsilon\vdots}$ for infinitely many n, then $\mathbb{P}\big(F_{n_0} \cap G\big) = 0$. From (3.88) we conclude that $\mathbb{P}(G) = 0$. $\qquad\square$

3.5 Multiple Point Sources

This section is devoted to proving Theorems 1.3.1 and 1.4.1.

3.5.1 Associativity and Hausdorff Continuity of the Smash Sum

In this section we establish two basic properties of the smash sum (1.6) that are needed to prove Theorems 1.3.1 and 1.4.1. In Lemma 3.5.1 we show that the smash sum is associative, and in Lemma 3.5.3 we show that it is continuous in the Hausdorff metric.

Lemma 3.5.1. *Let $A, B, C \subset \mathbb{R}^d$ be bounded open sets whose boundaries have measure zero. Then*

$$(A \oplus B) \oplus C = A \oplus (B \oplus C)$$
$$= A \cup B \cup C \cup D$$

where D is given by (3.10) with $\sigma = 1_A + 1_B + 1_C$.

Proof. Let

$$\gamma(x) = -|x|^2 - G(1_A + 1_B)(x)$$

and

$$\hat{\gamma}(x) = -|x|^2 - G(1_{A \oplus B} + 1_C)(x).$$

Let $u = s - \gamma$ and $\hat{u} = \hat{s} - \hat{\gamma}$, where $s, \hat{s}$ are the least superharmonic majorants of $\gamma, \hat{\gamma}$. Then

$$(A \oplus B) \oplus C = (A \oplus B) \cup C \cup \{\hat{u} > 0\}$$
$$= A \cup B \cup \{u > 0\} \cup C \cup \{\hat{u} > 0\}$$
$$= A \cup B \cup C \cup \{u + \hat{u} > 0\}. \tag{3.89}$$

Let ν_n be the final mass density for the divisible sandpile in $\delta_n \mathbb{Z}^d$ started from source density $1_{A^{::}} + 1_{B^{::}}$, and let u_n be the corresponding odometer function. By Theorem 3.2.1 we have $u_n \to u$ as $n \to \infty$. Moreover, by Theorem 3.2.7 we have

$$\nu_n^{\square}(x) \to 1_{A \oplus B}(x) \tag{3.90}$$

for all $x \notin \partial(A \oplus B)$. Let $\hat{u}_n$ be the odometer function for the divisible sandpile on $\delta_n \mathbb{Z}^d$ started from source density $\nu_n + 1_{C^{::}}$. By Proposition 3.1.13(i) the right side of (3.90) is continuous almost everywhere, so by Theorem 3.2.1 we have $\hat{u}_n \to \hat{u}$ as $n \to \infty$. On the other hand, by the abelian property, Lemma 2.2.1, the sum $u_n + \hat{u}_n$ is the odometer function for the divisible sandpile on $\delta_n \mathbb{Z}^d$ started from source density $1_{A^{::}} + 1_{B^{::}} + 1_{C^{::}}$, so by Theorem 3.2.1 we have $u_n + \hat{u}_n \to \tilde{u} := \tilde{s} - \tilde{\gamma}$, where

$$\tilde{\gamma}(x) = -|x|^2 - G(1_A + 1_B + 1_C)(x)$$

and $\tilde{s}$ is the least superharmonic majorant of $\tilde{\gamma}$. In particular,

$$u + \hat{u} = \lim_{n \to \infty} (u_n + \hat{u}_n) = \tilde{u},$$

so the right side of (3.89) is equal to $A \cup B \cup C \cup D$. $\qquad \square$

The following lemma shows that the smash sum of two sets with a small intersection cannot extend very far beyond their union. As usual, $\mathcal{L}$ denotes Lebesgue measure in $\mathbb{R}^d$, and A^ϵ denotes the outer ϵ-neighborhood of a set $A \subset \mathbb{R}^d$.

Lemma 3.5.2. *Let $A, B \subset \mathbb{R}^d$ be bounded open sets whose boundaries have measure zero, and let $\rho = (\mathcal{L}(A \cap B))^{1/d}$. There is a constant c, independent of A and B, such that*

$$A \oplus B \subset (A \cup B)^{c\rho}.$$

Proof. Let γ, s be given by (3.8) and (3.9) with $\sigma = 1_A + 1_B$, and write $u = s - \gamma$. Fix $x \in (A \oplus B) - (A \cup B)$ and let $r = \text{dist}(x, A \cup B)$. Let $\mathcal{B} = B(x, r/2)$. By Lemma 3.1.2(iii), s is harmonic in $A \oplus B - (A \cup B)$, so by Lemma 3.1.4 the function

$$\begin{aligned} w(y) &= u(y) - |x - y|^2 \\ &= s(y) + |y|^2 + G(1_A + 1_B)(y) - |x - y|^2 \end{aligned}$$

is harmonic on the intersection $(A \oplus B) \cap \mathcal{B}$; hence it attains its maximum on the boundary. Since $w(x) > 0$ the maximum cannot be attained on $\partial(A \oplus B)$, so it is attained at some point $y \in \partial\mathcal{B}$, and

$$u(y) \geq w(y) + \frac{r^2}{4} > \frac{r^2}{4}. \tag{3.91}$$

If z is any point outside $A \oplus B$, then by Lemma 3.1.6 and Lemma 3.1.19 with $\lambda = 4d$, there is a constant $c' \geq 1$ such that $u \leq c'h^2$ on $B(z, h)$ for all h. Taking $h = r/2\sqrt{c'}$, we conclude from (3.91) that $B(y, h) \subset A \oplus B$. Since $B(y, h)$ is disjoint from $A \cup B$, we have by Corollary 3.1.14

$$\mathcal{L}(A \cup B) + \omega_d h^d \leq \mathcal{L}(A \oplus B) = \mathcal{L}(A) + \mathcal{L}(B),$$

hence $\omega_d h^d \leq \mathcal{L}(A \cap B)$. Taking $c = 2\sqrt{c'}/\omega_d^{1/d}$ yields $r = c\omega_d^{1/d}h \leq c\rho$. $\qquad \square$

The following lemma, together with Lemma 3.1.8, shows that the smash sum operation is continuous in the Hausdorff metric.

Lemma 3.5.3. *Let $A, B \subset \mathbb{R}^d$ be bounded open sets whose boundaries have measure zero. For any $\epsilon > 0$ there exists $\eta > 0$ such that*

$$(A \oplus B)_\epsilon \subset A_\eta \oplus B_\eta \subset A^\eta \oplus B^\eta \subset (A \oplus B)^\epsilon. \tag{3.92}$$

Proof. By Lemma 3.1.20, since $A \oplus B = A \cup B \cup D$, we have

$$(A \oplus B)_\epsilon \subset A_{\epsilon'} \cup B_{\epsilon'} \cup D_{\epsilon'}$$

for some $\epsilon' > 0$. By Lemma 3.1.17(iii) with $\sigma = 1_A + 1_B$ and $\sigma_n = 1_{A_{1/n}} + 1_{B_{1/n}}$, for sufficiently small η we have $D_{\epsilon'} \subset A_\eta \oplus B_\eta$. This proves the first inclusion in (3.92).

The second inclusion is immediate from Lemma 3.1.8.

For the final inclusion, write $A' = A^\eta - A$ and $B' = B^\eta - B$. Since $\mathcal{L}(A') \downarrow \mathcal{L}(\partial A) = 0$ and $\mathcal{L}(B') \downarrow \mathcal{L}(\partial B) = 0$ as $\eta \downarrow 0$, for small enough η we have by Lemmas 3.5.1 and 3.5.2

$$A^\eta \oplus B^\eta = A \oplus B \oplus A' \oplus B'$$
$$\subset ((A \oplus B) \cup A^\eta \cup B^\eta)^{\epsilon - \eta}$$
$$\subset (A \oplus B)^\epsilon. \quad \square$$

3.5.2 Smash Sums of Balls

In this section we deduce Theorem 1.3.1 from our other results. The following result on internal DLA with a single point source is a restatement of the main result of [29].

Theorem 3.5.4. *Fix $\lambda > 0$, and let I_n be the random set of occupied sites for internal DLA in $\delta_n \mathbb{Z}^d$, starting with $m = \lfloor \lambda \delta_n^{-d} \rfloor$ particles at the origin. If $\delta_n \leq 1/n$ for all n, then for any $\epsilon > 0$, we have with probability one*

$$B_\epsilon^{::} \subset I_n \subset B^{\epsilon::} \qquad \text{for all sufficiently large } n,$$

where B is the ball of volume λ centered at the origin in $\mathbb{R}^d$.

Proof of Theorem 1.3.1. For $i = 1, \ldots, k$ let D_n^i, R_n^i, I_n^i be the domain of occupied sites in $\delta_n \mathbb{Z}^d$ starting from a single point source of $\lfloor \delta_n^{-d} \lambda_i \rfloor$ particles at $x_i^{::}$. By Theorems 1.2.1 and 1.2.2, for any $\eta > 0$ we have

$$(B_i)_\eta^{::} \subset D_n^i, R_n^i \subset (B_i)^{\eta::} \qquad \text{for all } i \text{ and all sufficiently large } n. \tag{3.93}$$

Moreover if $\delta_n \leq 1/n$, then by Theorem 3.5.4 we have with probability one

$$(B_i)_\eta^{::} \subset I_n^i \subset (B_i)^{\eta::} \qquad \text{for all } i \text{ and all sufficiently large } n. \tag{3.94}$$

Next we argue that the domains D_n, R_n, I_n can be understood as smash sums of D_n^i, R_n^i, I_n^i as i ranges over the integers $1, \ldots, k$. By the abelian property [15], the domain I_n is the Diaconis-Fulton smash sum in $\delta_n \mathbb{Z}^d$ of the domains I_n^i. Likewise, if r_1 is an arbitrary rotor configuration on $\delta_n \mathbb{Z}^d$, let S_n^2 be the smash sum of R_n^1 and R_n^2 formed using rotor-router dynamics with initial rotor configuration r_1, and let r_2 be the resulting final rotor configuration. For $i \geq 3$ define S_n^i inductively as the smash sum of S_n^{i-1} and R_n^i formed using rotor-router dynamics with initial rotor configuration r_{i-1}, and let r_i be the resulting final rotor configuration. Then $R_n = S_n^k$. Finally, by Lemma 2.2.1, the domain D_n contains the smash sum of domains D_n^i formed using divisible sandpile dynamics, and D_n is contained in the smash sum of the domains $D_n^i \cup \partial D_n^i$.

Fixing $\epsilon > 0$, by Theorem 1.1.3 and Lemma 3.1.8, it follows from (3.93) that for all sufficiently large n

$$A_{\epsilon/2}^{::} \subset D_n, R_n, I_n \subset \widetilde{A}^{\epsilon/2::} \tag{3.95}$$

where A is the smash sum of the balls $(B_i)_\eta$, and $\widetilde{A}$ is the smash sum of the balls $(B_i)^\eta$. By Lemma 3.5.3 we can take η sufficiently small so that

$$D_{\epsilon/2} \subset A \subset \widetilde{A} \subset D^{\epsilon/2}$$

where $D = B_1 \oplus \ldots \oplus B_k$. Together with (3.95), this completes the proof. $\qquad \square$

3.5.3 Algebraic Boundary

In this section we prove Theorem 1.4.1 showing that the boundary of a smash sum of disks in $\mathbb{R}^2$ is an algebraic curve. The first step is to establish the classical quadrature identity (1.12), which holds in any dimension. We will prove the following quadrature inequality generalizing (1.12) for superharmonic functions on a smash sum of balls in $\mathbb{R}^d$.

Proposition 3.5.5. *Fix $x_1, \ldots, x_k \in \mathbb{R}^d$ and $\lambda_1, \ldots, \lambda_k > 0$. Let B_i be the ball of volume λ_i centered at x_i. Then*

$$\int_{B_1 \oplus \ldots \oplus B_k} h(x) dx \leq \sum_{i=1}^{k} \lambda_i h(x_i)$$

for all integrable superharmonic functions h on $B_1 \oplus \ldots \oplus B_k$.

We could deduce Proposition 3.5.5 from Proposition 3.1.11 by integrating smooth approximations to h against smooth densities σ_n converging to the sum of the indicators of the B_i. However, it more convenient to use the following result of M. Sakai, which is proved by a similar type of approximation argument.

Theorem 3.5.6. *[42, Thm. 7.5] Let $\Omega \subset \mathbb{R}^d$ be a bounded open set, and let σ be a bounded function on $\mathbb{R}^d$ supported on $\bar{\Omega}$ satisfying $\sigma > 1$ on Ω. Let γ, s, D be given by (3.8)-(3.10).*

$$\int_D h(x)dx \le \int_D h(x)\sigma(x)dx$$

for all integrable superharmonic functions h on D.

We remark that the form of the obstacle in [42] is superficially different from ours: taking $w = 1$ and $\omega = \sigma$ in section 7 of [42], the obstacle is given by

$$\psi(x) = G1_B - G\sigma$$

for a large ball B, rather than our choice (3.8) of

$$\gamma(x) = -|x|^2 - G\sigma.$$

Note, however that $\gamma - \psi$ is constant on B by (3.6) and (3.7). Thus if B is sufficiently large, the two obstacle problems have the same noncoincidence set D by Lemmas 3.1.9 and 3.1.16.

To prove Proposition 3.5.5 using Theorem 3.5.6, we must produce an appropriate density σ on $\mathbb{R}^d$ strictly exceeding 1 on its support. We will take σ to be twice the sum of indicators of balls of half the volume of the B_i.

Proof of Proposition 3.5.5. Let B_i' be the ball of volume $\lambda_i/2$ centered at x_i, and consider the sum of indicators

$$\sigma = 2\sum_{i=1}^k 1_{B_i'}.$$

Let γ, s, D be given by (3.8)-(3.10). By Theorem 3.5.6 and the mean value proprerty for superharmonic functions (3.1), we have

$$\int_D h(x)dx \le 2\sum_{i=1}^k \int_{B_i'} h(x)dx \le \sum_{i=1}^k \lambda_i h(x_i)$$

for all integrable superharmonic functions h on D.

It remains to show that $D = B_1 \oplus \ldots \oplus B_k$. By Lemma 3.1.7 we have $B_i' \subset D$ for all i, hence by Lemma 3.5.1

$$D = B_1' \oplus B_1' \oplus \ldots \oplus B_k' \oplus B_k'.$$

By Lemma 3.1.15 we have $B_i = B_i' \oplus B_i'$, completing the proof. $\qquad\square$

Theorem 1.4.1 now follows from Proposition 3.5.5 by a result of Gustafsson [22, sec. 6]; see also [23, Lemma 1.1(a)]. The methods of [22] rely heavily on complex analysis, which is why they apply only in dimension two. Indeed, in higher dimensions it is not known whether classical quadrature domains have boundaries given by algebraic surfaces [44, Ch. 2].

Chapter 4

Aggregation on Trees

4.1 The Sandpile Group of a Tree

In this section we describe a short exact sequence relating the sandpile group of a tree to those of its principal subtrees. In the case of a regular tree this sequence splits, enabling us to compute the full decomposition of the sandpile group into cyclic subgroups. This resolves in the affirmative a conjecture of E. Toumpakari concerning the ranks of the Sylow p-subgroups.

We begin with a simple combinatorial problem. Let T_n be the d-regular tree of height n. Collapse all the leaves of T_n to a single vertex s, the *sink*, and add an edge connecting the root to the sink.

Lemma 4.1.1. *Let t_n be the number of oriented spanning trees of T_n rooted at the sink. Then*

$$t_n = t_{n-1}^{d-2}(dt_{n-1} - (d-1)t_{n-2}^{d-1}).$$

Proof. If the edge (r, s) from the root to the sink is included in the spanning tree, then each of the principal branches of T_n may be assigned an oriented spanning tree independently, so there are t_{n-1}^{d-1} such spanning trees. On the other hand, if (r, s) is not included in the spanning tree, there is a directed path $r \to x_1 \to \ldots \to x_{n-1} \to s$ in the spanning tree from the root to the sink. In this case, every principal branch except the one rooted at x_1 may be assigned an oriented spanning tree independently; within the branch rooted at x_1, every subbranch except the one rooted at x_2 may be assigned an oriented spanning tree independently; and so on (see Figure 4.1). Since there are $(d-1)^{n-1}$ possible paths
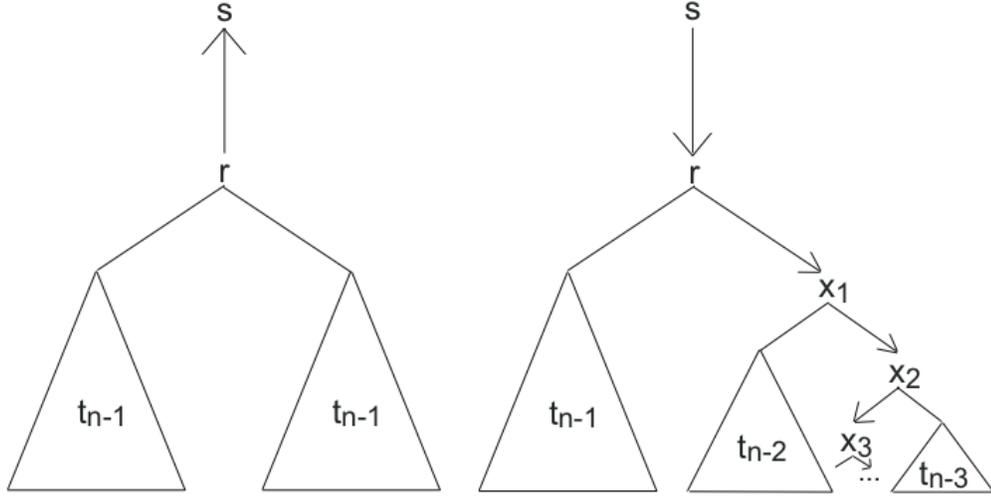
Figure 4.1: The two cases in the proof of Lemma 4.1.1

$x_1 \to \ldots \to x_{n-1}$, we conclude that

$$t_n = t_{n-1}^{d-1} + (d-1)^{n-1} \prod_{k=1}^{n-1} t_k^{d-2}. \tag{4.1}$$

Substituting $n-1$ for n we find that

$$(d-1)^{n-2} \prod_{k=1}^{n-2} t_k^{d-2} = t_{n-1} - t_{n-2}^{d-1}$$

hence from (4.1)

$$t_n = t_{n-1}^{d-2}(t_{n-1} + (d-1)(t_{n-1} - t_{n-2}^{d-1})). \quad \square$$

From Lemma 4.1.1 one can readily show by induction that

$$t_n = (1 + a + \ldots + a^{n-1}) \prod_{k=1}^{n-2} (1 + a + \ldots + a^k)^{a^{n-2-k}(a-1)}.$$

where $a = d - 1$. A variant of this formula was found by Toumpakari [47], who gives an algebraic proof.

For any graph G there is an abelian group, the *sandpile group*, whose order is the number of oriented spanning trees of G rooted at a fixed vertex; its definition and properties

are reviewed in section 4.1.1. A product formula such as the one above immediately raises the question of an analogous factorization of the sandpile group. Our main goal in this section is to prove Theorem 1.5.2, which establishes precisely such a factorization. In section 4.2 we will apply this result to study the rotor-router model on regular trees.

The next three sections are organized as follows. In section 4.1.1 we review the definition and basic properties of the sandpile group, and characterize the recurrent states on a tree explicitly in terms of what we call *critical vertices*. We prove a general result, Theorem 4.1.5, relating the sandpile group of an arbitrary tree to the sandpile groups of its principal branches (i.e. the subtrees rooted at the children of the root). This result takes the form of an isomorphism between two quotients. In section 4.1.2 we take advantage of the symmetry of regular trees to define a projection map back onto the subgroup being quotiented. This allows us to express $SP(T_n)$ as the direct sum of a cyclic group and a quotient of the direct sum of $d-1$ copies of $SP(T_{n-1})$, which enables us to prove Theorem 1.5.2 by induction. Finally, in section 4.1.3, we prove Toumpakari's conjecture [47].

4.1.1 General Trees

Let T be a finite rooted tree. Collapse all the leaves to a single vertex s, the *sink*, and add an edge connecting the root to the sink. If the vertices of T are $x_1, \ldots, x_n = s$, the *sandpile group* of T is defined by

$$SP(T) = \mathbb{Z}^n/\Delta$$

where $\Delta \subset \mathbb{Z}^n$ is the lattice

$$\Delta = \langle s, \Delta_{x_1}, \ldots, \Delta_{x_{n-1}} \rangle.$$

Here Δ_{x_i} is the vector in $\mathbb{Z}^n$ taking value 1 at each neighbor of x_i, value $-\deg(x)$ at x, and value 0 elsewhere. The sandpile group of a graph was originally defined in [5, 6, 14].

A nonnegative vector $u \in \mathbb{Z}^n$ may be thought of as a *chip configuration* with u_i chips at vertex x_i. A vertex $x \neq s$ is *unstable* if $u(x) \geq \deg(x)$. An unstable vertex may *topple*, sending one chip to each neighbor. Note that the operation of toppling the vertex x corresponds to adding the vector Δ_x to u. We say that u is *stable* if no vertex is unstable. Given chip configurations u and v, we define $u+v$ as the stable configuration resulting from starting with $u(x) + v(x)$ chips at each vertex x and toppling any unstable vertices; the order in which topplings are performed does not affect the final configuration, as observed by Dhar [14] and (in a more general setting) Diaconis and Fulton [15].

A chip configuration u is *recurrent* [13] if $u + v = u$ for a nonnegative configuration v. In [13] it is proved that every equivalence class mod Δ has a unique recurrent representative. Thus the sandpile group $SP(T)$ may be thought of as the set of recurrent configurations under the operation of addition followed by toppling. Note that if v is a nonnegative configuration, its recurrent representative is given by

$$\hat{v} := v + e$$

where e is the identity element of $SP(T)$ (the recurrent representative of 0); indeed, $\hat{v}$ is recurrent since e is recurrent, and $\hat{v} \equiv v \pmod{\Delta}$ since $e \in \Delta$. Note that if u is a recurrent configuration and v is a nonnegative configuration, then

$$u + \hat{v} = u + (v + e) = (u + e) + v = u + v. \tag{4.2}$$

The following result is a simple variant of the "burning algorithm" [14]; see also [13, Cor. 2.6].

Lemma 4.1.2. *Let $\beta(x)$ be the number of edges from x to the sink. A chip configuration on T is recurrent if and only if adding $\beta(x)$ chips at each vertex x causes every vertex to topple exactly once.*

Proof. Note that

$$\beta = \deg(s)s + \Delta_s = \deg(s)s - \sum_{x \neq s} \Delta_x. \tag{4.3}$$

If every vertex topples exactly once, then

$$u + \beta = u + \beta + \sum_{x \neq s} \Delta_x = u,$$

so u is recurrent. Conversely, suppose u is recurrent. Since $\beta \in \Delta$ we have $\hat{\beta} = e$, hence from (4.2)

$$u + \beta = u + e = u.$$

By (4.3), since $\{\Delta_x\}_{x \neq s}$ are linearly independent, every vertex topples exactly once. $\qquad \square$

We first characterize the recurrent configurations of a tree explicitly. The characterization uses the following inductive definition. Denote by $C(x)$ the set of children of a vertex $x \in T$.
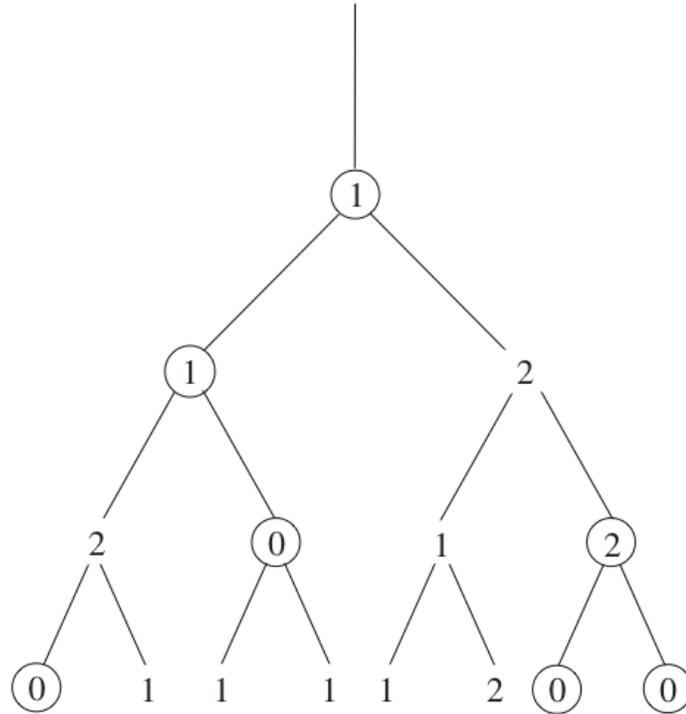
Figure 4.2: A recurrent configuration on the ternary tree of height 5. Critical vertices are circled; if any of the circled vertices had fewer chips, the configuration would not be recurrent.

Definition. A vertex $x \in T$ is *critical* for a chip configuration u if $x \neq s$ and

$$u(x) \leq \#\{y \in C(x) \mid y \text{ is critical}\}. \tag{4.4}$$

Proposition 4.1.3. *A configuration $u \in SP(T)$ is recurrent if and only if equality holds in (4.4) for every critical vertex x.*

Proof. If x is critical, then

$$u(x) + \#\{y \in C(x) \mid y \text{ is not critical}\} \leq \deg(x) - 1. \tag{4.5}$$

Thus after chips are added as prescribed by Lemma 4.1.2, inducting upward in decreasing distance to the root, if $x \neq r$ is critical, its parent must topple before it does. In particular, if strict inequality holds in (4.4), and hence in (4.5), for some vertex x, that vertex will never topple, so u is not recurrent.

Conversely, suppose equality holds in (4.4), hence in (4.5), for every critical x. Begin toppling vertices in order of decreasing distance from the root. Note that a non-

critical vertex x satisfies

$$u(x) + \#\{y \in C(x) \mid y \text{ is not critical}\} \geq \deg(x), \tag{4.6}$$

Inducting upward, every non-critical vertex topples once. Hence by equality in (4.5), once all vertices other than the root are stable, every critical vertex x has either toppled (if its parent toppled) or is left with exactly $\deg(x) - 1$ chips (if its parent did not topple). In particular, the root now topples, as it was given an extra chip in the beginning. Now if x is a critical vertex that has not yet toppled, its parent is also such a vertex. Inducting downward from the root, since all of these vertices are primed with $\deg(x) - 1$ chips, they each topple once, and u is recurrent. $\qquad\square$

Write $T_1, \ldots, T_k$ for the principal branches of T (i.e. the subtrees rooted at the children of the root). If u_i is a chip configuration on T_i, and a is an integer, we will use the notation $\begin{pmatrix} a \\ u_1, \ldots, u_k \end{pmatrix}$ for the configuration on T which has a chips at the root and coincides with u_i on T_i. The following result is an immediate consequence of Proposition 4.1.3.

Lemma 4.1.4. *Let* $u = \begin{pmatrix} a \\ u_1, \ldots, u_k \end{pmatrix}$.

(i) *If* u *is recurrent, each* u_i *is recurrent.*

(ii) *If* $u_1, \ldots, u_k$ *are recurrent and* $a = k$, *then* u *is recurrent.*

Write δ_x for a single chip at a vertex x, and denote by $\hat{x} = e + \delta_x$ the recurrent form of δ_x. Note that by (4.2), if u is recurrent then

$$u + \hat{x} = u + \delta_x. \tag{4.7}$$

Theorem 4.1.5. *Let* $T_1, \ldots, T_k$ *be the principal branches of* T. *Then*

$$SP(T)/(\hat{r}) \simeq \bigoplus_{i=1}^{k} SP(T_i)/((\hat{r}_1, \ldots, \hat{r}_k))$$

where r, r_i *are the roots of* T, T_i *respectively.*

Proof. Define $\phi : SP(T) \to \bigoplus_{i=1}^{k} SP(T_i)$ by

$$\begin{pmatrix} a \\ u_1, \ldots, u_k \end{pmatrix} \mapsto (u_1, \ldots, u_k). \tag{4.8}$$

Lemma 4.1.4(i) ensures this map is well-defined. Note that if

$$\begin{pmatrix} a \\ u_1, \ldots, u_k \end{pmatrix} = \begin{pmatrix} b \\ v_1, \ldots, v_k \end{pmatrix} + \hat{r},$$

by (4.7) either $b < k$ and $u_i = v_i$ for all i; or $b = k$ and the root topples, in which case $u_i = v_i + \hat{r}_i$. Thus ϕ descends to a map of quotients $\bar{\phi} : SP(T)/(\hat{r}) \to \bigoplus_{i=1}^k SP(T_i)/((\hat{r}_1, \ldots, \hat{r}_k))$.

By adding two configurations without allowing the root to topple, the configurations on each branch add independently, hence by (4.7)

$$\begin{pmatrix} a \\ u_1, \ldots, u_k \end{pmatrix} + \begin{pmatrix} b \\ v_1, \ldots, v_k \end{pmatrix} = \begin{pmatrix} c \\ u_1 + v_1, \ldots, u_k + v_k \end{pmatrix} + d\hat{r}$$

for some nonnegative integers c, d. Thus $\bar{\phi}$ is a group homomorphism. Moreover, $\bar{\phi}$ is surjective by Lemma 4.1.4(ii). Finally, if

$$(u_1, \ldots, u_k) = (v_1, \ldots, v_k) + c(\hat{r}_1, \ldots, \hat{r}_k),$$

then by (4.7), allowing the root to topple exactly c times, we obtain

$$\begin{pmatrix} a \\ u_1, \ldots, u_k \end{pmatrix} = \begin{pmatrix} b \\ v_1, \ldots, v_k \end{pmatrix} + (c(k+1) + d)\hat{r},$$

for a suitable integer d. Thus $\bar{\phi}$ is injective. $\qquad\square$

4.1.2   Regular Trees

In this section we show that for regular trees, Theorem 4.1.5 can be strengthened to express $SP(T)$ as a direct sum.

Let T_n be the regular tree of degree d and height n, with leaves collapsed to the sink vertex and an edge added from the root to the sink as in section 4.1.1. The chip configurations which are constant on the levels of T_n form a subgroup of $SP(T_n)$. If each vertex at height k has a_k chips, we can represent the configuration as a vector $(a_1, \ldots, a_{n-1})$. If such a recurrent configuration is zero on a level, all vertices above that level are critical, so by Proposition 4.1.3 they must have $d-1$ chips each. The recurrent configurations constant on levels are thus in bijection with integer vectors $(a_1, \ldots, a_{n-1})$ with $0 \le a_i \le d-1$ subject to the constraint that if $a_i = 0$ then $a_1 = \ldots = a_{i-1} = d - 1$.

$\hat{r}$	$2\hat{r}$	$3\hat{r}$	$4\hat{r}$	$5\hat{r}$	$6\hat{r}$	$7\hat{r}$	$8\hat{r}$	$9\hat{r}$	$10\hat{r}$	$11\hat{r}$	$12\hat{r}$	$13\hat{r}$	$14\hat{r}$	$15\hat{r}=e$
2	0	1	2	0	1	2	2	2	0	1	2	0	1	2
0	1	1	1	2	2	2	2	0	1	1	1	2	2	2
2	2	2	2	2	2	2	0	1	1	1	1	1	1	1

Figure 4.3: Multiples of the root $\hat{r}$ in the ternary tree of height 4. Each column vector represents a chip configuration which is constant on levels of the tree.

The following lemma uses the lexicographic order given by $a < b$ if for some k we have $a_{n-1} = b_{n-1}, \ldots, a_{k+1} = b_{k+1}$ and $a_k < b_k$. In the *cyclic lexicographic order* on recurrent vectors we have also $(d-1, \ldots, d-1) < (d-1, \ldots, d-1, 0)$.

Lemma 4.1.6. *If u, v are recurrent configurations on T_n that are constant on levels, write $u \rightsquigarrow v$ if v follows u in the cyclic lexicographic order on the set of recurrent vectors. Then for every integer $k \geq 0$, we have*

$$k\hat{r} \rightsquigarrow (k+1)\hat{r}.$$

Figure 4.3 demonstrates the lemma for a ternary tree of height 4.

Proof. By (4.7) we have

$$(k+1)\hat{r} = k\hat{r} + \delta_r.$$

Thus if $k\hat{r} = (a_1, \ldots, a_{n-1})$ with $a_1 < d-1$, then $(k+1)\hat{r} = (a_1+1, a_2, \ldots, a_{n-1})$ as desired. Otherwise, if not all a_i equal $d-1$, let $j > 1$ be such that $a_1 = \ldots = a_{j-1} = d-1$ and $a_j < d-1$. Adding a chip at the root initiates the toppling cascade

$$
\begin{pmatrix} d \\ d-1 \\ d-1 \\ \vdots \\ d-1 \\ d-1 \\ a_j \\ a_{j+1} \\ \vdots \\ a_{n-1} \end{pmatrix}
\rightarrow
\begin{pmatrix} 0 \\ d \\ d-1 \\ \vdots \\ d-1 \\ d-1 \\ a_j \\ a_{j+1} \\ \vdots \\ a_{n-1} \end{pmatrix}
\rightarrow
\begin{pmatrix} d-1 \\ 0 \\ d \\ \vdots \\ d-1 \\ d-1 \\ a_j \\ a_{j+1} \\ \vdots \\ a_{n-1} \end{pmatrix}
\rightarrow \ldots \rightarrow
\begin{pmatrix} d-1 \\ d-1 \\ d-1 \\ \vdots \\ 0 \\ d \\ a_j \\ a_{j+1} \\ \vdots \\ a_{n-1} \end{pmatrix}
\rightarrow
\begin{pmatrix} d-1 \\ d-1 \\ d-1 \\ \vdots \\ d-1 \\ 0 \\ a_j+1 \\ a_{j+1} \\ \vdots \\ a_{n-1} \end{pmatrix},
$$

as desired. If all $a_i = d-1$ the cascade will travel all the way down, ending in $(d-1, \ldots, d-1, 0)$ as desired. $\qquad\square$

Proposition 4.1.7. *Let T_n be the regular tree of degree d and height n, and let $R(T_n)$ be the subgroup of $SP(T_n)$ generated by $\hat{r}$. Then $R(T_n)$ consists of all recurrent configurations that are constant on levels, and its order is*

$$\#R(T_n) = \frac{(d-1)^n - 1}{d-2}. \tag{4.9}$$

Proof. Since the toppling rule is symmetric, all configurations in $R(T_n)$ are constant on levels. The number of such recurrent configurations is the number of vectors of the form $(d-1, \ldots, d-1, 0, a_j, \ldots, a_{n-1})$, with $a_i \in [d-1]$, which is

$$\sum_{j=0}^{n-1} (d-1)^j = \frac{(d-1)^n - 1}{d-2}.$$

Moreover, by Lemma 4.1.6, any such vector can be expressed as a multiple of $\hat{r}$, so $R(T_n)$ contains all the recurrent configurations that are constant on levels. $\qquad\square$

Index the vertices of the d-regular tree of height n by words of length $\leq n-2$ in the alphabet $\{1, \ldots, d-1\}$. Let σ_i be the automorphism of the tree given by

$$\sigma_i(w_1 \ldots w_k) = w_1 \ldots (w_i + 1) \ldots w_k$$

with the sum taken mod $d-1$; if $k < i$ then $\sigma_i(w) = w$. Given a map $\alpha : [n-2] \to [d-1]$ let σ_α be the composition $\prod_{i=1}^{n-2} \sigma_i^{\alpha(i)}$.

If σ is an automorphism of the form σ_α, write σu for the chip configuration $\sigma u(x) = u(\sigma x)$. Writing $u \oplus v$ for addition in the sandpile group and $u + v$ for the ordinary vector sum, we have

$$u \oplus v = u + v + \sum_{j=1}^{m} \Delta_{x_j}$$

Since $\sigma \Delta_x = \Delta_{\sigma x}$ we obtain

$$\sigma(u \oplus v) = \sigma(u) + \sigma(v) + \sum_{j=1}^{m} \Delta_{\sigma x_j}.$$

The configuration on the right side is stable, recurrent, and $\equiv \sigma(u) + \sigma(v) \pmod{\Delta}$, so it is equal to $\sigma(u) \oplus \sigma(v)$. Thus σ is an automorphism of the sandpile group.
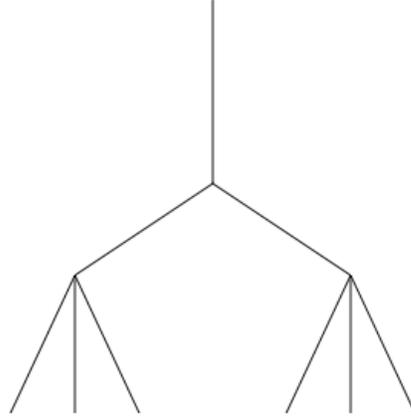
Figure 4.4: A non-regular tree for which Proposition 4.1.8 fails.

Proposition 4.1.8. *Let T_n be the regular tree of degree d and height n, and let $R(T_n) = (\hat{r})$ be the subgroup of $SP(T_n)$ generated by the root. Then*

$$SP(T_n) \simeq R(T_n) \oplus \frac{SP(T_{n-1}) \oplus \ldots \oplus SP(T_{n-1})}{(R(T_{n-1}), \ldots, R(T_{n-1}))}$$

with $d - 1$ summands of $SP(T_{n-1})$ on the right side.

Proof. Define $p : SP(T_n) \to SP(T_n)$ by

$$p(u) = (d-1)^2 \sum_{\alpha : [n-2] \to [d-1]} \sigma_\alpha u. \tag{4.10}$$

By construction $p(u)$ is constant on levels, so the image of p lies in $R(T_n)$ by Proposition 4.1.7. Given $u \in R(T_n)$, since u is constant on levels we have $\sigma_\alpha(u) = u$ for all α. Since there are $(d-1)^{n-2}$ terms in the sum (4.10), we obtain

$$p(u) = (d-1)^n u = u$$

where the second inequality follows from (4.9). Thus $R(T_n)$ is a summand of $SP(T_n)$, and the result follows from Theorem 4.1.5. $\qquad\square$

Proposition 4.1.8 fails for general trees. For example, if T is the tree consisting of a root with 2 children each of which have 3 children (Figure 4.4), then $\hat{r} = \begin{pmatrix} 2 \\ 3,3 \end{pmatrix}$ has order

10 and the element $x = \begin{pmatrix} 2 \\ 0,3 \end{pmatrix}$ satisfies $4x = \hat{r}$, so x has order 40. The total number of recurrent configurations is $4 \cdot 4 \cdot 3 - 8 = 40$, so $SP(T) \simeq \mathbb{Z}/40\mathbb{Z}$, and $R(T) \simeq \mathbb{Z}/10\mathbb{Z}$ is not a summand.

Write $\mathbb{Z}_p^q$ as a shorthand for $(\mathbb{Z}/p\mathbb{Z}) \oplus \ldots \oplus (\mathbb{Z}/p\mathbb{Z})$ with q summands.

Theorem 4.1.9. *Let T_n be the regular tree of degree $d = a + 1$ and height n, with leaves collapsed to the sink vertex and an edge joining the root to the sink. Then*

$$SP(T_n) \simeq \mathbb{Z}_{1+a}^{a^{n-3}(a-1)} \oplus \mathbb{Z}_{1+a+a^2}^{a^{n-4}(a-1)} \oplus \ldots \oplus \mathbb{Z}_{1+a+\ldots+a^{n-2}}^{a-1} \oplus \mathbb{Z}_{1+a+\ldots+a^{n-1}}.$$

Proof. Induct on n. For the base case $n = 2$ we have

$$SP(T_2) = R(T_2) \simeq \mathbb{Z}_d = \mathbb{Z}_{1+a}.$$

Write $q_n = 1 + a + \ldots + a^{n-1}$. By Proposition 4.1.7, the root subgroup $R(T_n)$ is cyclic of order q_n. By Proposition 4.1.8 and the inductive hypothesis, it follows that

$$SP(T_n) \simeq \mathbb{Z}_{q_n} \oplus \frac{SP(T_{n-1})^{\oplus a}}{\mathbb{Z}_{q_{n-1}}}$$

$$\simeq \mathbb{Z}_{q_n} \oplus \mathbb{Z}_{q_{n-1}}^{a-1} \oplus \mathbb{Z}_{q_{n-2}}^{a(a-1)} \oplus \ldots \oplus \mathbb{Z}_{q_2}^{a^{n-3}(a-1)}. \quad \square$$

4.1.3 Proof of Toumpakari's Conjecture

As before write $a = d - 1$ and

$$q_n = 1 + a + \ldots + a^{n-1}.$$

If p is a prime not dividing $d(d-1)$, let t_p be the least positive n for which $p | q_n$. Then

$$t_p = \begin{cases} p & \text{if } a \equiv 1 \ (\text{mod } p) \\ \text{ord}_p(a), & \text{else.} \end{cases}$$

Here $\text{ord}_p(a)$ is the least positive k for which $p | a^k - 1$. Note that $p | q_n$ if and only if $t_p | n$. The following result was conjectured by E. Toumpakari in [47] (where the factor of $d - 2$ was left out, presumably an oversight).

Theorem 4.1.10. *Let B_n be the ball of radius n in the d-regular tree, with each leaf connected by $d-1$ edges to the sink, but with no edge connecting the root to the sink. Let p be a prime not dividing $d(d-1)$, and let $S_p(n)$ be the Sylow-p subgroup of the sandpile group $SP(B_n)$. Then*

$$\operatorname{rank}(S_p(n)) = \begin{cases} d(d-2) \sum_{\substack{m<n \\ m\equiv n\,(\operatorname{mod}\,t_p)}} (d-1)^m, & \text{if } n \not\equiv -1 \ (\operatorname{mod}\,t_p); \\ d(d-2) \sum_{\substack{m<n \\ m\equiv n\,(\operatorname{mod}\,t_p)}} (d-1)^m + d - 1, & \text{if } n \equiv -1(\operatorname{mod}\,t_p). \end{cases}$$

Proof. By Theorem 4.1.5 we have

$$SP(B_n)/(\hat{r}) \simeq \frac{SP(T_{n+1}) \oplus \ldots \oplus SP(T_{n+1})}{(R(T_{n+1}), \ldots, R(T_{n+1}))}$$

with d summands. By Proposition 4.1.7 we have $R(T_{n+1}) \simeq \mathbb{Z}_{q_{n+1}}$, so from Theorem 4.1.9

$$SP(B_n)/(\hat{r}) \simeq \mathbb{Z}_{q_{n+1}}^a \oplus \mathbb{Z}_{q_n}^{(a-1)(a+1)} \oplus \mathbb{Z}_{q_{n-1}}^{(a-1)a(a+1)} \oplus \ldots \oplus \mathbb{Z}_{q_2}^{(a-1)a^{n-2}(a+1)}. \qquad (4.11)$$

By Proposition 7.2 of [47], the root subgroup $(\hat{r})$ of $SP(B_n)$ has order $d(d-1)^n$. Thus for p not dividing $d(d-1)$ the Sylow p-subgroup of $SP(B_n)$ is the same as that of the quotient $SP(B_n)/(\hat{r})$. Each summand $\mathbb{Z}_{q_k}$ in (4.11) contributes 1 to the rank of $S_p(n)$ if $t_p|k$ and 0 otherwise. If $n \not\equiv -1 \ (\operatorname{mod}\,t_p)$, the total rank is therefore

$$\operatorname{rank}(S_p(n)) = \sum_{\substack{2\leq k\leq n \\ t_p|k}} (a-1)a^{n-k}(a+1)$$

$$= d(d-2) \sum_{\substack{0\leq m\leq n-2 \\ m\equiv n\,(\operatorname{mod}\,t_p)}} (d-1)^m.$$

In the case that $n \equiv -1$ (mod t_p), the first summand $\mathbb{Z}_{q_{n+1}}^a$ in (4.11) contributes an additional rank $a = d-1$ to $S_p(n)$. $\qquad \square$

4.2 The Rotor-Router Model on Trees

This section is devoted to proving Theorems 1.5.1 and 1.5.3.

4.2.1 The Rotor-Router Group

In this section we define the *rotor-router group* of a graph and show it is isomorphic to the sandpile group. This isomorphism, Theorem 4.2.5, is mentioned in the physics

literature; see [38, 37]. To our knowledge the details of the proof are not written down anywhere. While our main focus is on the tree, the isomorphism is just as easily proved for general graphs, and it seems to us worthwhile to record the general proof here.

Let G be a strongly connected finite directed graph without loops. Fix a sink vertex s in G, and write $Rec(G)$ for the set of oriented spanning trees of G rooted at the sink. Given a configuration of rotors T, write $e_x(T)$ for the configuration resulting from starting a chip at x and letting it walk according to the rotor-router rule until it reaches the sink. (Note that if the chip visits a vertex infinitely often, it visits all of its neighbors infinitely often; since G is strongly connected, the chip eventually reaches the sink.) We view T as a subgraph of G in which every vertex except the sink has out-degree one. Note that such a subgraph is an oriented spanning tree rooted at the sink if and only if it has no oriented cycles.

Lemma 4.2.1. *If $T \in Rec(G)$, then $e_x(T) \in Rec(G)$.*

Proof. Let Y be any collection of vertices of G. If the chip started at x reaches the sink without ever visiting Y, then the rotors at vertices in Y point the same way in $e_x(T)$ as they do in T, so they do not form an oriented cycle. If the chip does visit Y, let $y \in Y$ be the last vertex it visits. Then either $y = s$, or the rotor at y points to a vertex not in Y; in either case, the rotors at vertices in Y do not form an oriented cycle. $\square$

We will need slightly more refined information about the intermediate states that occur before the chip falls into the sink. These states may contain oriented cycles, but only of a very restricted form. For a vertex x we write $Cyc_x(G)$ for the set of rotor configurations U such that

(i) U contains an oriented cycle; and

(ii) If the rotor $U(x)$ is deleted, the resulting configuration contains no oriented cycles.

Lemma 4.2.2. *Starting from a rotor configuration $T_0 \in Rec(G)$ with a chip at x_0, let T_k and x_k be the rotor configuration and chip location after k steps. Then*

(i) *If $T_k \notin Rec(G)$, then $T_k \in Cyc_{x_k}(G)$.*

(ii) *If $T_k \in Rec(G)$, then $x_k \notin \{x_0, \ldots, x_{k-1}\}$.*

Proof. (i) It suffices to show that any oriented cycle in T_k contains x_k. Let Y be any set of vertices of G not containing x_k. If Y is disjoint from $\{x_0, \ldots, x_{k-1}\}$, then the rotors at vertices in Y point the same way in T_k as they do in T_0, so they do not form an oriented cycle. Otherwise, let $y \in Y$ be the vertex visited latest before time k. The rotor at y points to a vertex not in Y, so the rotors at vertices in Y do not form an oriented cycle.

(ii) Suppose $x_k \in \{x_0, \ldots, x_{k-1}\}$. Let $y_0 = x_k$, and for $i = 0, 1, \ldots$ let $y_{i+1} = T_k(y_i)$. Then the last exit from x_k before time k was to y_1, and by induction if $y_1, \ldots, y_{i-1}$ are disjoint from x_k, then y_{i-1} was visited before time k, and the last exit from y_{i-1} before time k was to y_i. It follows that $y_i = x_k$ for some i, and hence T_k contains an oriented cycle. □

Lemma 4.2.3. *If $T_1, T_2 \in Rec(G)$ and $e_x(T_1) = e_x(T_2)$, then $T_1 = T_2$.*

Proof. We will show that T can be recovered from $e_x(T)$ by reversing one rotor step at a time. Given rotor configurations U, U' and vertices y, y', we say that (U', y') is a predecessor of (U, y) if a chip at y' with rotors configured according to U' would move to y in a single step with resulting rotors configured according to U. For each neighbor $z \to y$ with $U(z) = y$ there is a unique predecessor of the form (U', z), which we will denote $P_z(U, y)$.

Suppose (U, y) is an intermediate state in the evolution from T to $e_x(T)$. If $U \notin Rec(G)$, then by case (i) of Lemma 4.2.2 there is a cycle of rotors $U(y) = y_1, U(y_1) = y_2, \ldots, U(y_n) = y$. If $U(z) = y$ and $z \neq y_n$, then z is not in this cycle, so the predecessor $P_z(U, y)$ has a cycle disjoint from its chip location. Thus $P_z(U, y)$ does not belong to $Rec(G)$ or to $Cyc_z(G)$, so by Lemma 4.2.2 it cannot be an intermediate state in the evolution from T to $e_x(T)$. The state immediately preceding (U, y) in the evolution from T to $e_x(T)$ must therefore be $P_{y_n}(U, y)$.

Now suppose $U \in Rec(G)$. By case (ii) of Lemma 4.2.2, U is the rotor configuration when y is first visited. If $y = x$, then $U = T$. Otherwise, let $x = x_0 \to x_1 \to \ldots \to x_k = s$ be the path in U from x to the sink. Then the last exit from x before visiting y was to x_1. By induction, if $x_1, \ldots, x_{j-1}$ are different from y, then x_{j-1} was visited before y and the last exit from x_{j-1} before visiting y was to x_j. It follows that $x_j = y$ for some j, and the state immediately preceding (U, y) must be $P_{x_{j-1}}(U, y)$. □

Thus for any vertex x of G, the operation e_x of adding a chip at x and routing it to the sink acts invertibly on the set of states $Rec(G)$ whose rotors form oriented spanning

trees rooted at the sink. It is for this reason that we call these states recurrent. We define the *rotor-router group* $RR(G)$ as the subgroup of the permutation group of $Rec(G)$ generated by $\{e_x\}_{x\in G}$. Note that if there are two (indistinguishable) chips on G and each takes a single step according to the rotor-router rule, the resulting rotor configuration does not depend on the order of the two steps. Thus the operators e_x commute, and the group $RR(G)$ is abelian; for a general discussion of this property, which is shared by a number of models including the abelian sandpile and the rotor-router, see [15].

Lemma 4.2.4. $RR(G)$ *acts transitively on* $Rec(G)$.

Proof. Given $T_1, T_2 \in Rec(G)$, let $u(x)$ be the number of rotor turns needed to get from $T_1(x)$ to $T_2(x)$. Let $v(x)$ be the number of chips ending up at x if $u(y)$ chips start at each vertex y, with rotors starting in configuration T_1, and each chip takes a single step. After each chip has taken a step, the rotors are in configuration T_2, hence

$$\left(\sum_{x\in V(G)} u(x)e_x\right) T_1 = \left(\sum_{x\in V(G)} v(x)e_x\right) T_2.$$

Letting $g = \sum_{x\in V(G)}(u(x) - v(x))e_x$ we obtain $T_2 = gT_1$. $\qquad\square$

Theorem 4.2.5. *Let G be a strongly connected finite directed graph without loops, let $RR(G)$ be its rotor-router group, and $SP(G)$ its sandpile group. Then $RR(G) \simeq SP(G)$.*

Proof. Let V be the vertex set of G. Recall [13] that the sandpile group can be expressed as the quotient

$$SP(G) \simeq \mathbb{Z}^V /(s, \Delta_x)_{x\in V}$$

where $s \in V$ is the sink and

$$\Delta_x = \sum_{y\leftarrow x} y - \text{outdeg}(x)x.$$

Define $\phi : \mathbb{Z}^G \to RR(G)$ by

$$\phi(u) = \sum_{x\in V(G)} u(x)e_x.$$

Since adding $\text{outdeg}(x)$ chips at x and letting each chip take one step results in one chip at each neighbor $y \leftarrow x$ with the rotors unchanged, we have

$$\text{outdeg}(x)e_x = \sum_{y\leftarrow x} e_y.$$

Thus $\phi(\Delta_x) = 0$ and ϕ descends to a map $SP(G) \to RR(G)$. This map is trivially surjective; to show it is injective, by Lemma 4.2.4 we have

$$\#RR(G) \geq \#Rec(G) = \#SP(G),$$

where the equality on the right is the matrix-tree theorem [46]. □

4.2.2 Aggregation on Regular Trees

Let T be the infinite d-regular tree. Fix an origin vertex o in T. In *rotor-router aggregation*, we grow a cluster of points in T by repeatedly starting chips at the origin and letting them walk until they exit the cluster. Beginning with $A_1 = \{o\}$, define the cluster A_n inductively by

$$A_n = A_{n-1} \cup \{x_n\}, \qquad n > 1.$$

where $x_n \in T$ is the endpoint of a rotor-router walk started at o and stopped on first exiting A_{n-1}. We do not change the positions of the rotors when adding a new chip.

In this section we use the group isomorphism proved in the last section to show that A_n is a perfect ball for suitable values of n (Theorem 1.5.1). The proof makes crucial use of the calculation of the order of the subgroup of the sandpile group of T_n generated by the root, Proposition 4.1.7.

A function H on the vertices of a directed graph G is *harmonic* if

$$H(x) = \frac{1}{\text{outdeg}(x)} \sum_{y \leftarrow x} H(y)$$

for all vertices x.

Lemma 4.2.6. *Let H be a harmonic function on the vertices of G. Suppose chips on G can be routed, starting with $u(x)$ chips at each vertex x and ending with $v(x)$ chips at each vertex x, in such a way that the initial and final rotor configurations are the same. Then*

$$\sum_{x \in V(G)} H(x)u(x) = \sum_{x \in V(G)} H(x)v(x).$$

Proof. Let $u = u_0, u_1, \ldots, u_k = v$ be the intermediate configurations. If u_{i+1} is obtained from u_i by routing a chip from x_i to y_i, then

$$\sum_{x \in V(G)} H(x)(u(x) - v(x)) = \sum_i H(x_i) - H(y_i). \tag{4.12}$$

If the initial and final rotor configurations are the same, then each rotor makes an integer number of full turns, so the sum in (4.12) can be written

$$\sum_i H(x_i) - H(y_i) = \sum_{x \in V(G)} N(x) \sum_{y \leftarrow x} (H(x) - H(y))$$

where $N(x)$ is the number of full turns made by the rotor at x. Since H is harmonic, the inner sum on the right vanishes. □

Let T_n be the regular tree of degree d and height n, with an edge added from the root r to the sink o. Denote by $(X_t)_{t \geq 0}$ the simple random walk on T_n, and let τ be the first hitting time of the set consisting of the leaves and the sink. Fix a leaf z of T_n, and let

$$H(x) = \mathbb{P}_x(X_\tau = z) \tag{4.13}$$

be the probability that random walk started at x and stopped at time τ stops at z.

The quantity $H(r)$ can be computed by a standard martingale argument. Recall that the process

$$M_t = a^{-|X_t|}$$

is a martingale, where $a = d - 1$ and $|x|$ denotes the distance from x to the sink. Since M_t has bounded increments and $\mathbb{E}_r \tau < \infty$, we obtain from optional stopping

$$a^{-1} = \mathbb{E}_r M_0 = \mathbb{E}_r M_\tau = p + (1 - p)a^{-n}$$

where $p = \mathbb{P}_r(X_\tau = o)$. Solving for p we obtain

$$\mathbb{P}_r(X_\tau = o) = \frac{a^{n-1} - 1}{a^n - 1}. \tag{4.14}$$

In the event that the walk stops at a leaf, by symmetry it is equally likely to stop at any leaf. Since there are a^{n-1} leaves, we obtain from (4.14)

$$H(r) = \frac{1 - \mathbb{P}_r(X_\tau = o)}{a^{n-1}} = \frac{a - 1}{a^n - 1}. \tag{4.15}$$

Lemma 4.2.7. *Let $a = d - 1$. If the initial rotor configuration on T_n is acyclic, then starting with $\frac{a^n - 1}{a - 1}$ chips at the root, and stopping each chip when it reaches a leaf or the sink, exactly one chip stops at each leaf, and the remaining $\frac{a^{n-1} - 1}{a - 1}$ chips stop at the sink. Moreover the starting and ending rotor configurations are identical.*

Proof. By Theorem 4.2.5 and Proposition 4.1.7, the element $e_r \in RR(T_n)$ has order $\frac{a^n-1}{a-1}$, so the starting and ending rotor configurations are identical. Fix a leaf z and let H be the harmonic function given by (4.13). By Lemma 4.2.6 and (4.15), the number of chips stopping at z is

$$\sum H(x)v(x) = \sum H(x)u(x) = \frac{a^n - 1}{a - 1}H(r) = 1.$$

Since there are a^{n-1} leaves, the remaining $\frac{a^n-1}{a-1} - a^{n-1} = \frac{a^{n-1}-1}{a-1}$ chips stop at the sink. $\quad\square$

Let T be the infinite d-regular tree. The ball of radius ρ centered at the origin in $o \in T$ is

$$B_\rho = \{x \in T \ : \ |x| \leq \rho\}$$

where $|x|$ is the length of the shortest path from o to x. Write

$$b_\rho = \#B_\rho = 1 + (a+1)\frac{a^\rho - 1}{a - 1}.$$

As the following result shows, provided we start with an acyclic configuration of rotors, the rotor-router aggregation cluster A_n is a perfect ball at those times when an appropriate number of chips have aggregated. It follows that at all other times, the cluster is as close as possible to a ball: if $b_\rho < n < b_{\rho+1}$ then $B_\rho \subset A_n \subset B_{\rho+1}$.

Theorem 4.2.8. *Let A_n be the region formed by rotor-router aggregation on the infinite d-regular tree, starting from n chips at the origin. If the initial rotor configuration is acyclic, then $A_{b_\rho} = B_\rho$ for all $\rho \geq 0$.*

Proof. Define a modified aggregation process A'_n as follows. Stop the n-th chip when it either exits the occupied cluster A'_{n-1} or returns to o, and let

$$A'_n = A'_{n-1} \cup \{x'_n\}$$

where x'_n is the point where the n-th chip stops. By relabeling the chips, this yields a time change of the original process, i.e. $A'_n = A_{f(n)}$ for some sequence $f(1), f(2), \ldots$. Thus it suffices to show $A'_{c_\rho} = B_\rho$ for some sequence $c_1, c_2, \ldots$. We will show by induction on ρ that this is the case for

$$c_\rho = 1 + (a+1)\sum_{t=1}^{\rho}\frac{a^t - 1}{a - 1},$$

and that after c_ρ chips have stopped the rotors are in their initial state. For the base case $\rho = 1$, we have $c_1 = a + 2 = d + 1$. The first chip stops at o, and the next d stop at each
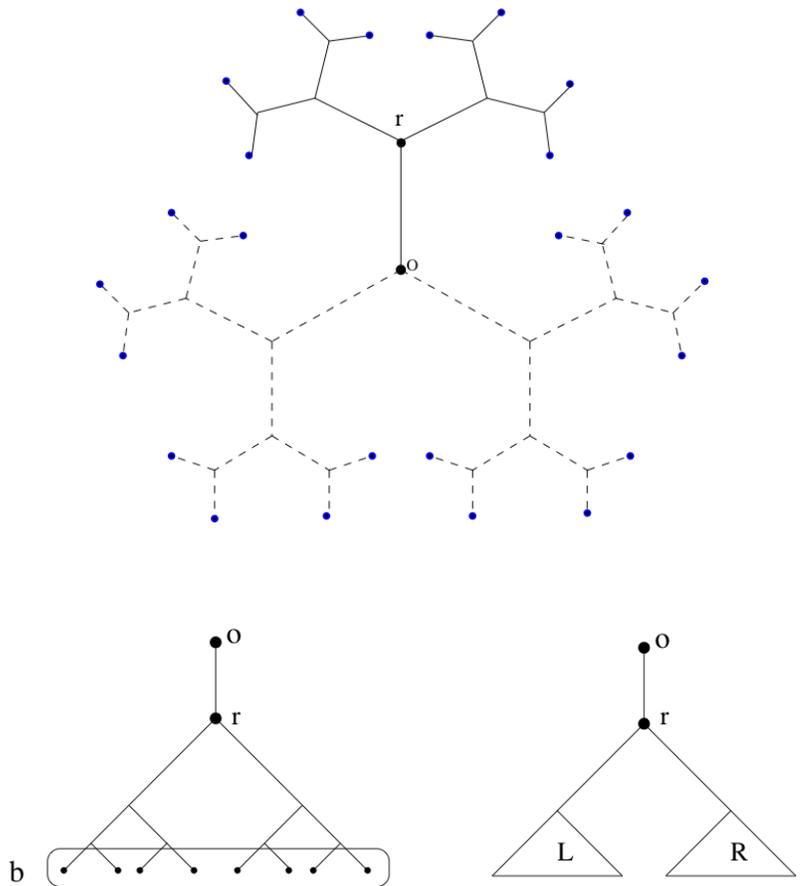
Figure 4.5: A branch of the regular ternary tree.

of the neighbors of o, so $A'_{d+1} = B_1$. Since the rotor at o has performed one full turn, it is back in its initial state.

Assume now that $A'_{c_{\rho-1}} = B_{\rho-1}$ and that the rotors are in their initial acyclic state. Starting with $c_\rho - c_{\rho-1}$ chips at o and letting each take one step, there are $\frac{a^\rho - 1}{a - 1}$ chips at each neighbor of o. Since the rotor at o has performed an integer number of full turns, the rotors remain in their initial acyclic state. By Lemma 4.2.7, exactly one chip will stop at each leaf $z \in B_\rho - B_{\rho-1}$, and the remainder will stop at o. Thus $A'_{c_\rho} = B_\rho$. Moreover, by Lemma 4.2.7, once all chips have stopped, the rotors are once again in their initial state, completing the inductive step. $\qquad\square$

4.2.3 Recurrence and Transience

In this section we explore questions of recurrence and transience for the rotor-router walk on regular trees. We aim to study to what extent the rotor-router walk behaves as a deterministic analogue of random walk. We find that the behavior depends quite dramatically on the initial configuration of rotors.

Let T_n be the d-regular tree of height n. We collapse the leaves to a sink labeled b for boundary, label the origin o and treat it as both the source and a sink. We will examine the hitting rates of the two sinks for various configurations of the model, and compare to the expected result for random walk.

Since the origin is a sink, the recurrence and transience of chips on one principal branch of the tree are independent of the rotor configurations on the other branches. Therefore, as shown in Figure 4.5, we focus on a single branch Y_n of T_n. By the root r of Y_n we will mean the unique node with $|r| = 1$. To each rotor direction we associate an index from $\{1, \ldots, d\}$, with direction d corresponding to a rotor pointing to the parent vertex. Rotors cycle through the indices in order.

Lemma 4.2.9. *Let Y_n be a principal branch of T_n, the regular ternary tree of height n. If all rotors initially point in direction 1, then the first $2^n - 1$ chips started at r alternate, the first stopping at b, the next stopping at o, the next at b, and so on until all the rotors again point in direction 1.*

Proof. Induct on n. For $n = 2$ the result is obvious as there is only one rotor, which sends the first chip in direction 2 to b, the next chip up in direction 3 to o, and the third chip in direction 1 to b, at which point the rotor is again in its initial state.

Now suppose that the lemma holds for Y_{n-1}. Let L and R be the two principal branches of Y_n. We think of L and R as each having a rotor that points either to b or back up to r. The initial state of these rotors is pointing to r. The first chip is sent from the origin to R, which by induction sends it to b. Note that the root rotor is now pointing towards R, the R-rotor is pointing to b, and the L-rotor is pointing to r (Figure 4.6a). We now observe that the next four chips form a pattern that will be repeated. The second chip is sent directly to o (Figure 4.6b) and the third chip is sent to L which sends it to b (Figure 4.6c). The fourth chip is sent to R, but by induction this chip is returned and then it is sent to o (Figure 4.6d). Finally, the fifth chip is sent to L, returned, sent to R, and through to b (Figure 4.6e). Note that the root rotor is now again pointing towards R, the
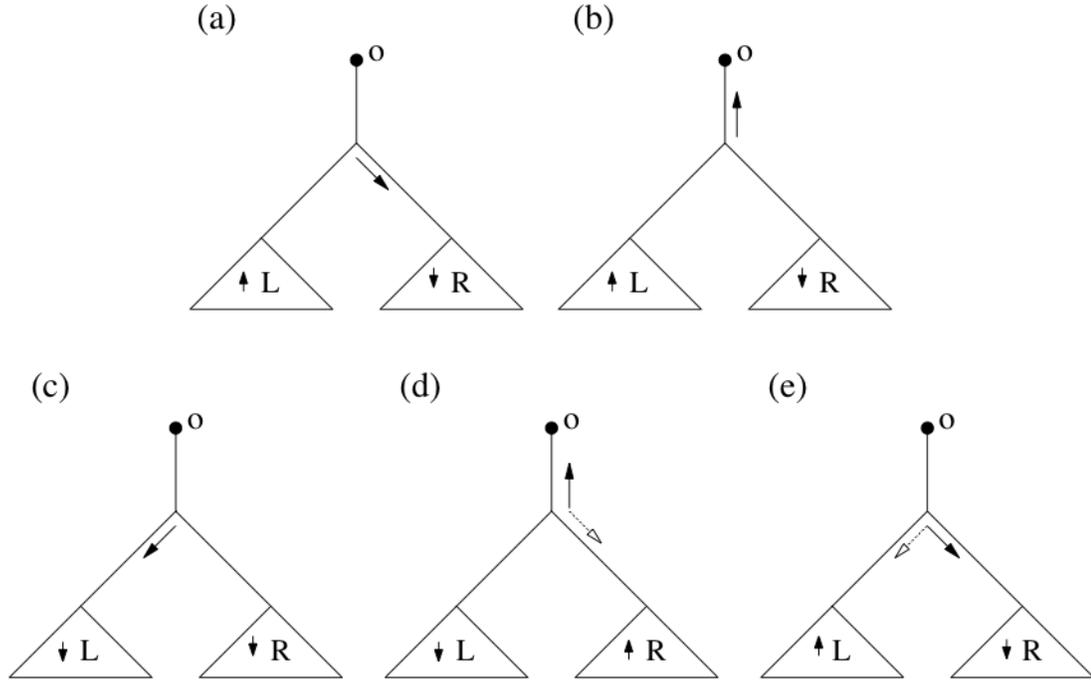
Figure 4.6: The four-chip cycle, which begins after the first chip has been routed to b.

R-rotor is again pointing to b, and the L-rotor is again pointing to r. In this cycle of four chips, the two branches R and L see two chips apiece. This cycle repeats $2^{n-2} - 1$ times, and each subtree sees $2^{n-1} - 2$ chips.

Recall that we sent the first chip to R, so that it has seen a total of $2^{n-1} - 1$ chips. By induction, all the rotors in R are in their initial configuration. We have sent a total of $2^n - 3$ chips. The next chip is sent to o, and the last to L which sends it to b. Now L has seen $2^{n-1} - 1$ chips so by induction all of its rotors are in their initial configuration. The root rotor is pointing towards L, its initial configuration. We have sent a total of $2^n - 1$ chips, alternating between b and o, and all of the rotors of Y_n are in the initial configuration, so the inductive step is complete. $\square$

We remark that the obvious generalization of Lemma 4.2.9 to trees of degree $d > 3$ fails; indeed, we do not know of a starting rotor configuration on trees of higher degree which results in a single chip stopping at o alternating with a string of $d - 1$ chips stopping at b.

Consider now the case of the infinite ternary tree T. A chip performing rotor-router walk on T must either return to o or escape to infinity visiting each vertex only

finitely many times. Thus the state of the rotors after a chip has escaped to infinity is well-defined. Therefore, we can properly define $R(m)$ as the number of chips that stop at o after m chips have executed the rotor-router walk beginning at r. The following result shows that there is a configuration of the tree for which the rotor-router walk behaves as an exact quasirandom analogue to the random walk, in which chips exactly alternate returning to the origin with escaping to infinity.

Proposition 4.2.10. *Let T be the infinite ternary tree, with principal branches labeled $Y^{(1)}$, $Y^{(2)}$, and $Y^{(3)}$ in correspondence with the direction indexing of the rotor at the origin. Set the rotors along the rightmost branch of $Y^{(3)}$ initially pointing in direction 2, and all remaining rotors initially pointing in direction 1. Let $E(m)$ be the expected number of chips that return to the origin if m chips perform independent random walks on T. Let $R(m)$ be the number of chips that return to the origin if m chips perform rotor-router walks on T. Then $|E(m) - R(m)| \leq \frac{1}{2}$ for all m.*

Proof. Lemma 4.2.9 implies that for the branches $Y^{(1)}$ and $Y^{(2)}$, the chips sent to a given branch alternate indefinitely with the first escaping to infinity, the next returning to o, and so on. Likewise, chips sent to $Y^{(3)}$ will alternate indefinitely with the first returning to o, the next escaping to infinity, and so on. Since chips on the full tree T are routed cyclically through the branches beginning with $Y^{(2)}$, we see that the chips too will alternate indefinitely between escaping to infinity and returning to the origin, with the first escaping to infinity. Thus $R(m) = \lfloor \frac{m}{2} \rfloor$. Taking $n \to \infty$ in (4.14) we obtain $E(m) = \frac{m}{2}$, and the result follows. $\qquad \square$

Lemma 4.2.11. *Let Y_n be a principal branch of T_n, the d-regular tree of height n. If all rotors initially point in direction $d - 1$, then the first $n - 1$ chips return to o before hitting height n of Y_n.*

Proof. Induct on n. For $n = 2$, there is a single rotor which sends the first chip from r to o. Now suppose the lemma holds for Y_{n-1}. Let $Z_1, \ldots, Z_{d-1}$ be the principal branches of Y_n. The first chip placed at r is sent directly to o. By the inductive hypothesis, the first $n - 2$ chips that are sent to each branch Z_i are returned to r before hitting height n of Y_n. Thus each of the next $n - 2$ chips is sent to Z_1, returned to r, sent to Z_2, and so on until it is sent to Z_{d-1}, returned to r and then routed to o. $\qquad \square$

Our next result shows that, perhaps surprisingly, the rotors can be set up so as to make rotor-router walk on the d-regular tree recurrent.

Proposition 4.2.12. *Let T be the infinite d-regular tree. If all rotors initially point in direction $d - 1$, then every chip in an infinite succession of chips started at the origin eventually returns to the origin.*

Proof. By Lemma 4.2.11, for each n, the n-th chip sent to each principal branch Y returns to the origin before hitting height $n + 1$ of T. $\square$

We continue our exploration of recurrence and transience on the infinite ternary tree T, but now we allow for arbitrary rotor configurations. We again focus on a single branch Y. For a given rotor configuration we define the *escape sequence* for the first n chips to be the binary word $a = a_1 \ldots a_n$, where for each j,

$$a_j = \begin{cases} 0, & \text{if the } j^{th} \text{ chip returns to the origin;} \\ 1, & \text{if the } j^{th} \text{ chip escapes to infinity.} \end{cases}$$

As noted above, a chip cannot stay within a finite height indefinitely without returning to the origin, so a is well-defined.

We define a map ψ on an escape sequence $a = a_1 \ldots a_n$. First we rewrite a as the concatenation of subwords $b_1 \cdots b_m$ where each $b_j \in \{0, 10, 110\}$. Since at least one of any three consecutive chips entering Y is routed directly upwards by the root rotor, at most two of any three consecutive letters in an escape sequence a can be ones. Therefore, any escape sequence can be factored in this way up to the possible concatenation of an extra 0. Now we define $\psi(a) = (c, d)$ by

$$(c_j, d_j) = \begin{cases} (0,0), & \text{if } b_j = 0 \\ (1,1), & \text{if } b_j = 110 \\ (0,1), & \text{if } b_j = 10 \text{ and } \#\{i < j | b_i = 10\} \text{ is odd} \\ (1,0), & \text{if } b_j = 10 \text{ and } \#\{i < j | b_i = 10\} \text{ is even} \end{cases}$$

Now we define the map ϕ on a pair of binary words c and d each of length m by

$\phi(c, d) = b_1 \cdots b_m$, where

$$b_j = \begin{cases} 0, & \text{if } (c_j, d_j) = (0, 0) \\ 10, & \text{if } (c_j, d_j) = (1, 0) \text{ or } (0, 1) \\ 110, & \text{if } (c_j, d_j) = (1, 1) \end{cases}$$

Note that ϕ is a left inverse of ψ, i.e. $\phi \circ \psi(a) = a$, up to possible concatenation of an extra 0.

Lemma 4.2.13. *Let Y be a principal branch of the infinite ternary tree. Fix a rotor configuration on Y with the root rotor pointing up. Let c and d be the escape sequences for the configurations on the left and right sub-branches of Y, respectively. Then $\phi(c, d)$ is the escape sequence for the full branch Y.*

Proof. We claim that each word b_j is the escape sequence for the jth full rotation of the root rotor. Suppose first that $(c_j, d_j) = (0, 0)$ so that $b_j = 0$. In this case, since each sub-branch will return the next chip it sees back to r, the next chip that enters the full tree will be sent up to the origin, and the root rotor will once again be pointing up to the origin. If $(c_j, d_j) = (1, 0)$, the next chip entering the full tree will be routed to the left sub-branch where it escapes to infinity. The following chip will be routed to the right sub-branch and then back up to the origin, and the root rotor will once again be pointing up. In this case we have $b_j = 10$. If $(c_j, d_j) = (0, 1)$, the next chip entering the full tree will be routed to the left sub-branch, back up to r, and then to the right sub-branch where it escapes to infinity. The following chip will be routed directly up to the origin leaving the root rotor pointed up once again. Again, in this case $b_j = 10$. Finally, if $(c_j, d_j) = (1, 1)$, the next two chips entering the full branch will escape to infinity, one through each sub-branch. The following chip will be routed directly up to the origin, once again leaving the root rotor pointing up. In this case we have $b_j = 110$. $\qquad\square$

We observe that if the root rotor is not pointing up and c and d are the escape sequences of the left and right sub-branches respectively, then $a = \phi(c, d)$ is *not* the escape sequence of the full branch. In order to calculate the escape sequence of the full branch we define *extended escape sequences* on the sub-branches, c' and d'. Suppose the root rotor initially points to the left sub-branch. Then $c' = 0c$ and $d' = d$. Suppose the root rotor initially points to the right sub-branch. Then $c' = 0c$ and $d' = 0d$. Now $a = \phi(c', d')$ is the escape sequence of the full branch.

We now introduce the condition that is central to characterizing which words can be escape sequences:

$$\text{any subword of length } 2^k - 1 \text{ contains at most } 2^{k-1} \text{ ones} \qquad (P_k)$$

We next show that the map ψ preserves this requirement.

Lemma 4.2.14. *Let a be a binary word satisfying (P_k) and let $\psi(a) = (c, d)$ as defined above. Then c and d each satisfy (P_{k-1}).*

Proof. Let c' be a subword of c of length $2^{k-1} - 1$ and let d' be the corresponding subword of d. Let $a' = \phi(c', d')$, which is a subword of $a0$. The formula for ϕ guarantees that a' has one zero for each letter of c', so a' has $2^{k-1} - 1$ zeros. Since the last letter of a' is zero, and a satisfies (P_k), it follows that a' has at most 2^{k-1} ones (else after truncating the final zero, the suffix of a' of length $2^k - 1$ has at most $2^{k-1} - 2$ zeros, hence at least $2^{k-1} + 1$ ones).

Let m be the number of ones in c'. Since the instances of $(0, 1)$ and $(1, 0)$ alternate in the formula for $\psi(a) = (c, d)$, it follows that d' must have at least $m - 1$ ones. Since the number of ones in c' and d' combined equals the number of ones in a', we obtain $2m - 1 \leq 2^{k-1}$, hence $m \leq 2^{k-2}$. The same argument with the roles of c and d reversed shows that d has at most 2^{k-2} ones. $\square$

Lemma 4.2.15. *Let $a = a_1 \ldots a_n$ be a binary word of length n. Then a is an escape sequence for some rotor configuration on the infinite branch Y if and only if a satisfies (P_k) for all k.*

Proof. Suppose a is an escape sequence. We prove that a satisfies (P_k) for each k by induction on k. That a satisfies (P_1) is trivial. Now suppose that every escape sequence satisfies (P_{k-1}) and let c and d be the extended escape sequences of the left and right sub-branches respectively. As we saw above, $a = \phi(c, d)$ up to the possible concatenation of an extra zero. Let a' be a subword of a of length $2^k - 1$, and let $\psi(a') = (c', d')$. Then there are words c'' and d'' each of which is a subword of c or d, and which are equal to c' and d', respectively, except possibly in the first letter; moreover the first letters satisfy $c'_1 \leq c''_1$ and $d'_1 \leq d''_1$.

By the formula for ψ, the number of ones in a' is the sum of the number of ones in c' and d'. If c' has length at most $2^{k-1} - 1$, then since c and d satisfy (P_{k-1}), each of c' and d' has at most 2^{k-2} ones, and therefore a' has at most 2^{k-1} ones. On the other hand,

if c' has length at least 2^{k-1}, then the number of zeros in a' is at least $2^{k-1} - 1$. Thus a' has at most 2^{k-1} ones, so a satisfies (P_k).

The proof of the converse is by induction on n. For $n = 1$ the statement is trivial. Suppose that every binary word of length $n - 1$ satisfying (P_k) for each k is an escape sequence. Then by Lemma 4.2.14, $\psi(a) = (c, d)$ gives a pair of binary words each satisfying (P_k) for all k. If c and d have length $n - 1$ or less, then they are escape sequences by induction, hence a is an escape sequence by Lemma 4.2.13. If c and d are of length n then the definition of ψ implies that $a_j = 0$ for all j, in which case a is an escape sequence by Proposition 4.2.12. $\square$

We can now establish our main result characterizing all possible escape sequences on the infinite ternary tree.

Theorem 4.2.16. *Let* $a = a_1 \ldots a_n$ *be a binary word. For* $j \in \{1, 2, 3\}$ *write* $a^{(j)} = a_j a_{j+3} a_{j+6} \ldots$. *Then* a *is an escape sequence for some rotor configuration on the infinite ternary tree* T *if and only if each* $a^{(j)}$ *satisfies* (P_k) *for all* k.

Proof. Let $Y^{(1)}$, $Y^{(2)}$, and $Y^{(3)}$ be the three principal branches of T assigned so that the rotor at the origin initially points to $Y^{(3)}$. Then a is the escape sequence for T if and only if $a^{(j)} = a_j a_{j+3} a_{j+6} \ldots$ is the escape sequence for $Y^{(j)}$. The result now follows from Lemma 4.2.15. $\square$

Chapter 5

Conjectures and Open Problems

5.1   Rotor-Router Aggregation

A number of intriguing questions remain unanswered about rotor-router aggregation in $\mathbb{Z}^d$. While we have shown in Theorem 1.2.1 that the asymptotic shape of the rotor-router model with a single point source of particles is a ball, the near perfect circularity found in Figure 1.3 remains a mystery. In particular, we do not know whether an analogue of Theorem 1.2.2 holds for the rotor-router model, with constant error in the radius as the number of particles grows.

Equally mysterious are the patterns in the rotor directions evident in Figure 1.3. The rotor directions can be viewed as values of the odometer function mod $2d$, but our control of the odometer is not fine enough to provide useful information about the patterns. If the rescaled occupied region $\sqrt{\pi/n}\,A_n$ is viewed as a subset of the complex plane, it appears that the monochromatic regions visible in Figure 1.3, in which all rotors point in the same direction, occur near points of the form $(1+2z)^{-1/2}$, where $z = a+bi$ is a Gaussian integer (i.e. $a, b \in \mathbb{Z}$). We do not even have a heuristic explanation for this phenomenon. Figure 5.1 shows the image of $A_{1,000,000}$ under the map $z \mapsto 1/z^2$; the monochromatic patches in the transformed region occur at lattice points.

László Lovász (personal communication) has asked whether the occupied region A_n is simply connected, i.e. whether its complement is connected. While Theorem 1.2.1 shows that A_n cannot have any holes far from the boundary, we cannot answer his question at present.

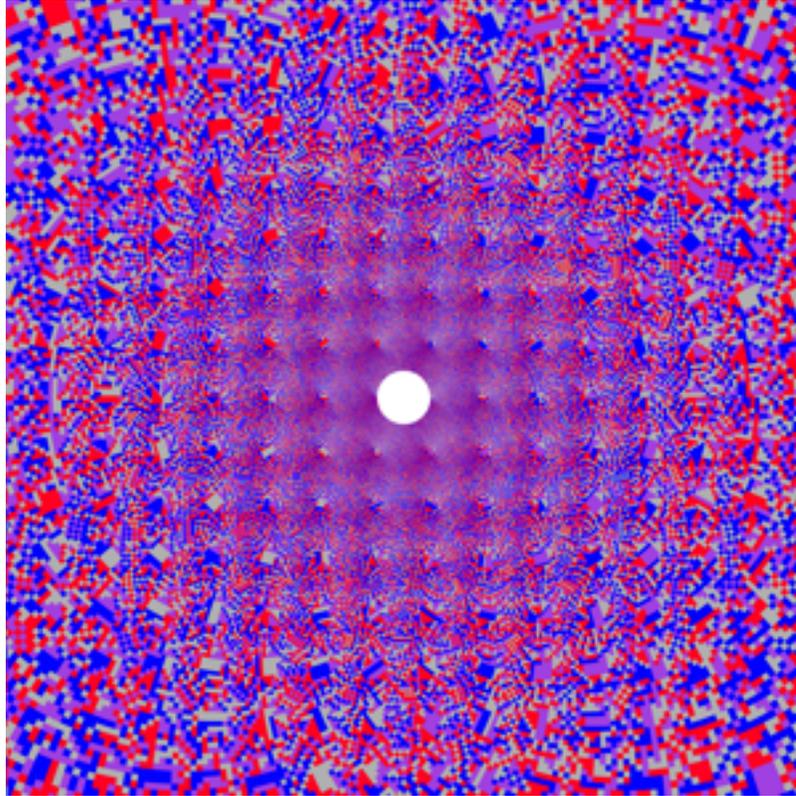Another question is whether our methods could be adapted to internal DLA to

Figure 5.1: Image of the rotor-router aggregate of one million particles under the map $z \mapsto 1/z^2$. The colors represent the rotor directions. The white disc in the center is the image of the complement of the occupied region.

show that if $n = \omega_d r^d$, then with high probability $B_{r-c \log r} \subset I_n$, where I_n is the internal DLA cluster of n particles. The current best bound is due to Lawler [30], who proves that with high probability $B_{r-r^{1/3}(\log r)^2} \subset I_n$.

Finally, let us mention two simple variants of aggregation models which appear to have interesting limit shapes, but about which we cannot prove anything at present. The following conjecture addresses a variant of the rotor-router model. Analogous variants can also be defined for internal DLA and the divisible sandpile, and by analogy with Theorem 1.1.2 we expect all three models to yield the same limiting shape.

Conjecture. Let A_n be the region formed from rotor-router aggregation starting with n particles at the origin in $\mathbb{Z}^2$, if rotors on the positive x-axis reflect particles back in the direction they came from; all other rotors operate normally. Then as $n \to \infty$, the asymptotic shape of the rescaled region $n^{-1/2} A_n$ is the cardioid in $\mathbb{R}^2$ whose boundary is given by the
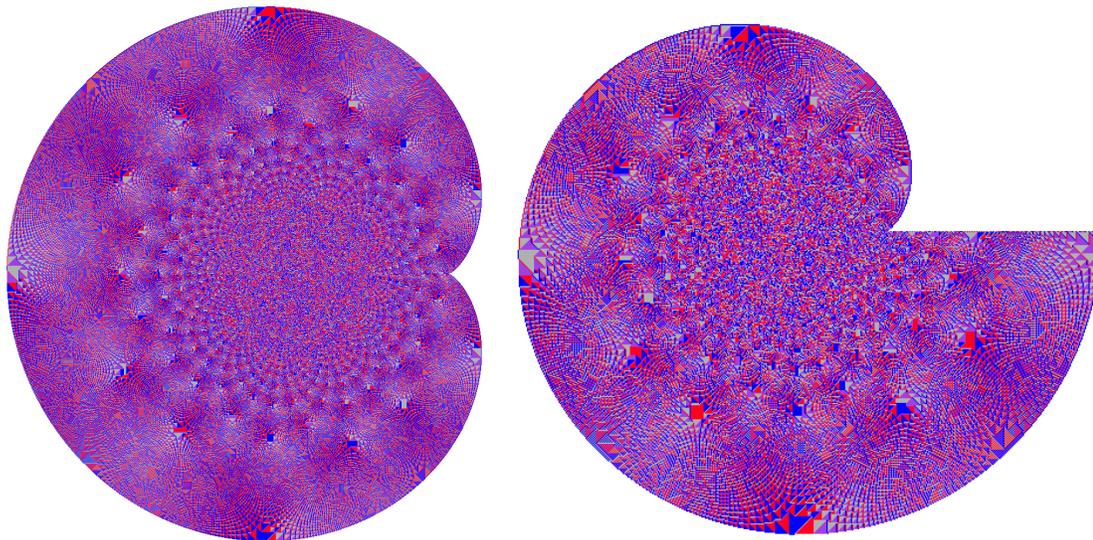
Figure 5.2: Rotor-router aggregation started from a point source at the origin in $\mathbb{Z}^2$, modified so that sites on the positive x-axis (a) reflect particles back in the direction they came from; or, (b) send all particles in the downward direction.

curve

$$\pi^2 |x|^4 - \pi |x|^2 + \frac{2\sqrt{2\pi}}{3\sqrt{3}} x_1 - \frac{1}{12} = 0.$$

This apparent cardioid is pictured in Figure 5.2 on the left. If instead we alter the rotors on the positive x-axis so that they always point downward, we obtain the spiral-shaped region on the right. Our simulations indicate that a limiting shape exists for this variant as well, but we do not have a conjectured formula for the boundary curve.

5.2 Sandpile Aggregation

While Theorem 1.2.3 improves on the best-known bounds of [31] and [18] for the shape of sandpile aggregation in $\mathbb{Z}^d$, it does not settle the question of existence of a limiting shape. There is also an intriguing pattern in the shapes of the sandpile aggregates $S_{n,H}$ of n particles in which every site starts with a hole of depth H. These shapes are pictured in Figure 5.3 for $n = 250,000$ and three different values of H.

Question. In dimension two, is the limiting shape $n^{-1/2} S_{n,H}$ a regular polygon with $4H+12$ sides?
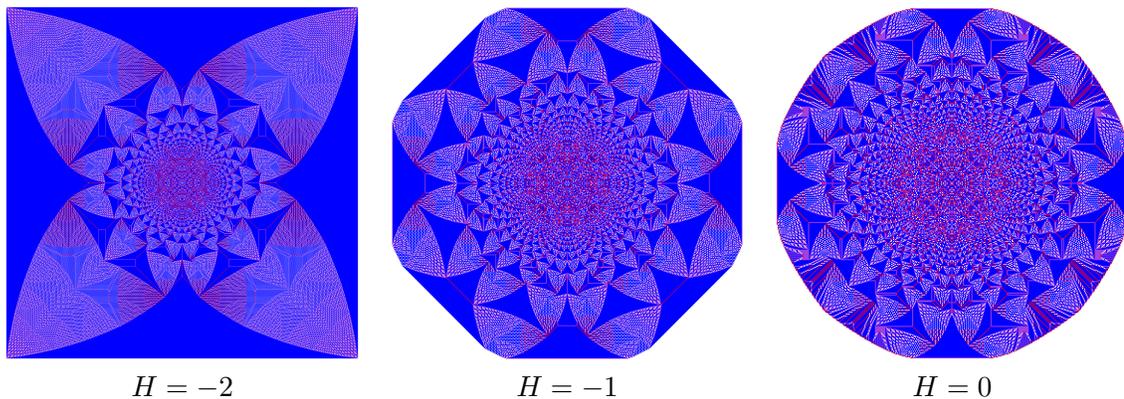
Figure 5.3: Abelian sandpile shape $S_{n,H}$ started from a point source of $n = 250,000$ particles in $\mathbb{Z}^2$, for three different values of the hole depth H.

Simulations indicate a regular polygon with some rounding at the corners; it remains unclear if the rounded portions of the boundary become negligible in the limit. If the limiting shape is not a polygon, it would still be very interesting to establish the weaker statement that it has the dihedral symmetry D_{4H+12}.

The only case that is even partly solved is $H = -2$: Fey and Redig [18] prove that $S_{n,-2}$ is a cube in $\mathbb{Z}^d$, but the growth rate of the cube is still not known. It would be of interest to show that the volume of the cube $S_{n,-2}$ is of order n.

Bibliography

[1] D. Aharonov and H. S. Shapiro, Domains on which analytic functions satisfy quadrature identities, *J. Analyse Math.* **30** (1976), 39–73.

[2] N. Alon and J. H. Spencer, *The Probabilistic Method*, John Wiley & Sons, 1992.

[3] S. Axler, P. Bourdon and W. Ramey, *Harmonic Function Theory*, 2nd ed., Springer, 2001.

[4] P. Bak, C. Tang and K. Wiesenfeld, Self-organized criticality: an explanation of the $1/f$ noise, *Phys. Rev. Lett.* **59**, no. 4 (1987), 381–384.

[5] N. L. Biggs, Chip-firing and the critical group of a graph. *J. Algebraic Combin.* **9** (1999), no. 1, 25–45.

[6] A. Björner, L. Lóvász and P. Shor, Chip-firing games on graphs, *European J. Combin.* **12** (1991), no. 4, 283–291.

[7] L. A. Caffarelli, The obstacle problem revisited, *J. Fourier Anal. Appl.* **4** (1998), no. 4-5, 383–402.

[8] L. A. Caffarelli, L. Karp and H. Shahgholian, Regularity of a free boundary problem with application to the Pompeiu problem, *Ann. Math.* **151** (2000), 269–292.

[9] D. Crowdy, "Quadrature Domains and Fluid Dynamics," in Quadrature Domains and Their Applications, *Operator Theory: Advances and Applications* **156** (2005), 113–129.

[10] W. Chen and T. Schedler, Concrete and abstract structure of the sandpile group for thick trees with loops, `http://arxiv.org/abs/math.CO/0701381`.

[11] M. Cook, The tent metaphor, available at
    `http://paradise.caltech.edu/∼cook/Warehouse/ForPropp/`.

[12] J. N. Cooper and J. Spencer, Simulating a random walk with constant error, *Combin. Probab. Comput.*, to appear; `http://www.arxiv.org/abs/math.CO/0402323`.

[13] R. Cori and D. Rossin, On the sandpile group of a graph, *European J. Combin.* **21** (2000), no. 4, 447–459.

[14] D. Dhar, Self-organized critical state of sandpile automaton models, *Phys. Rev. Lett.* **64** (1990), 1613–1616.

[15] 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, 95–119.

[16] J. L. Doob, *Classical Potential Theory and Its Probabilistic Counterpart*, Springer, 1984.

[17] L. C. Evans, *Partial Differential Equations*, AMS Graduate Studies in Mathematics, 1998.

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

[19] A. Friedman, *Variational Principles and Free-Boundary Problems*, John Wiley & Sons, 1982.

[20] Y. Fukai and K. Uchiyama, Potential kernel for two-dimensional random walk. *Ann. Probab.* **24** (1996), no. 4, 1979–1992.

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

[22] B. Gustafsson, Quadrature Identities and the Schottky double, *Acta Appl. Math.* **1** (1983), 209–240.

[23] B. Gustafsson, Singular and special points on quadrature domains from an algebraic geometric point of view, *J. Analyse Math.* **51** (1988), 91–117.

[24] B. Gustafsson and H. S. Shapiro, "What is a quadrature domain?" in Quadrature Domains and Their Applications, *Operator Theory: Advances and Applications* **156** (2005), 1–25.

[25] L. Karp and H. Shahgholian, Regularity of a free boundary problem, *J. Geom. Anal.* **9** (1999), 653–669.

[26] M. Kleber, Goldbug variations, *Math. Intelligencer* **27** (2005), no. 1, 55–63.

[27] P. Koosis, La plus petite majorante surharmonique et son rapport avec l'existence des fonctions entières de type exponentiel jouant le rôle de multiplicateurs, *Ann. Inst. Fourier (Grenoble)* **33** (1983), fasc. 1, 67–107.

[28] G. Lawler, *Intersections of Random Walks*, Birkhäuser, 1996.

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

[30] G. Lawler, Subdiffusive fluctuations for internal diffusion limited aggregation, *Ann. Probab.* **23** (1995) no. 1, 71–86.

[31] Y. Le Borgne and D. Rossin, On the identity of the sandpile group, *Discrete Math.* **256** (2002) 775–790.

[32] L. Levine, The rotor-router model, Harvard University senior thesis (2002), `http://arxiv.org/abs/math/0409407`.

[33] L. Levine and Y. Peres, Spherical asymptotics for the rotor-router model in $\mathbb{Z}^d$, *Indiana Univ. Math. Journal*, to appear; `http://arxiv.org/abs/math/0503251`.

[34] L. Levine and Y. Peres, The rotor-router shape is spherical, *Math. Intelligencer* **27** (2005), no. 3, 9–11.

[35] E. H. Lieb and M. Loss, *Analysis*, 2nd ed., AMS Graduate Studies in Mathematics, vol. 14, 2001.

[36] T. Lindvall, *Lectures on the Coupling Method*, Wiley, 1992.

[37] A. M. Povolotsky, V. B. Priezzhev, R. R. Shcherbakov, Dynamics of Eulerian walkers, *Phys. Rev.* E **58** (1998), 5449–54.

[38] V. B. Priezzhev, D. Dhar, A. Dhar, and S. Krishnamurthy, Eulerian walkers as a model of self-organised criticality, *Phys. Rev. Lett.* **77** (1996) 5079–82.

[39] J. Propp, Three lectures on quasirandomness, available at `http://www.math.wisc.edu/`∼`propp/berkeley.html`.

[40] S. Richardson, Hele-Shaw flows with a free boundary produced by the injection of fluid into a narrow channel, *J. Fluid Mech.* **56** (1972), 609–618.

[41] M. Sakai, *Quadrature Domains*, Lect. Notes Math. **934**, Springer, 1982.

[42] M. Sakai, Solutions to the obstacle problem as Green potentials, J. Analyse Math. **44** (1984/85), 97–116.

[43] H. Shahgholian, On quadrature domains and the Schwarz potential, *J. Math. Anal. Appl.* **171** (1992), 61–78.

[44] H. S. Shapiro, *The Schwarz Function and its Generalization to Higher Dimensions*, Univ. of Arkansas Lect. Notes Math., vol. 9, Wiley, 1992.

[45] F. Spitzer, *Principles of Random Walk*, Springer, 1976.

[46] R. P. Stanley, *Enumerative Combinatorics*, vol. 2, Cambridge University Press, 1999.

[47] E. Toumpakari, On the sandpile group of regular trees, *European J. Combin.* **28**, no. 3 (2007), 822–842.

[48] K. Uchiyama, Green's functions for random walks on $\mathbb{Z}^N$, *Proc. London Math. Soc.* **77** (1998), no. 1, 215–240.

[49] J. Van den Heuvel, Algorithmic aspects of a chip-firing game, *Combin. Probab. Comput.* **10**, no. 6 (2001), 505–529.

[50] A. N. Varchenko and P. I Etingof, *Why the Boundary of a Round Drop Becomes a Curve of Order Four*, AMS University Lecture Series, vol. 3, 1992.