Random Spatial Structures and Sums

Stochastische ruimtelijke structuren en sommen

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit
Utrecht, op gezag van de Rector Magnificus,
Prof. dr. W.H. Gispen, ingevolge het besluit van het College voor
Promoties in het openbaar te verdedigen op
donderdag 13 juni 2002 des ochtends te 10:30 uur

door

Lorna Booth

geboren op 23 december 1976, te Stockport, Engeland

Promotor : Prof. Dr. R.W.J. Meester
Faculteit der Exacte Wetenschappen,
Vrije Universiteit Amsterdam.

Random Spatial Structures and Sums
Lorna Booth
Faculteit Wiskunde en Informatica, Universiteit Utrecht
Proefschrift Universiteit Utrecht – met samenvatting in het Nederlands

ISBN: 90-393-3052-2

Mathematics Subject Classification (2000): 60D05, 60K35, 82B20, 82B21, 82B26, 82B41, 82B43, 82C22, 94C99.

Contents

Chapter 1

Introduction

This thesis contains articles about certain random structures in infinite space. It can be difficult to think about such large systems, but many of the ideas in this thesis are quite easy to explain, and play with. For this reason the introduction is interactive, with puzzles and problems for the reader, which I hope will give some insight into the principles underlying the articles. Some answers are given at the end of the introduction. Chapters 2, 3 and 4 are the articles themselves.

1.1 Sum puzzles

Look at Figure 1.1. You see a grid of circles, with a $+1$ or -1 inside each of them. I made this picture by flipping a fair coin for each circle. If the coin came up heads I wrote a $+1$, if tails a -1.

There is a light grey path through the grid, from the (S) in the bottom left hand corner to the (F) in the top right. We can follow this path, from the (S), and add up the +1s and -1s as we go. We find the sums

$$\begin{aligned} &+1, \\ &+2 = (+1) + (+1), \\ &+1 = (+1) + (+1) + (-1), \\ &+2 = (+1) + (+1) + (-1) + (+1), \text{ etc.} \end{aligned}$$

The list of all the sums is,

$$\begin{aligned} &+1, +2, +1, +2, +1, 0, -1, -2, -3, -4, -5, -4, -5, -4, -3, \\ &-4, -3, -4, -3. - 4, -5, -4, -3. \end{aligned}$$

All of these sums are between -5 and $+2$. We say that all of these sums lie in the *interval* $[-5, +2]$.

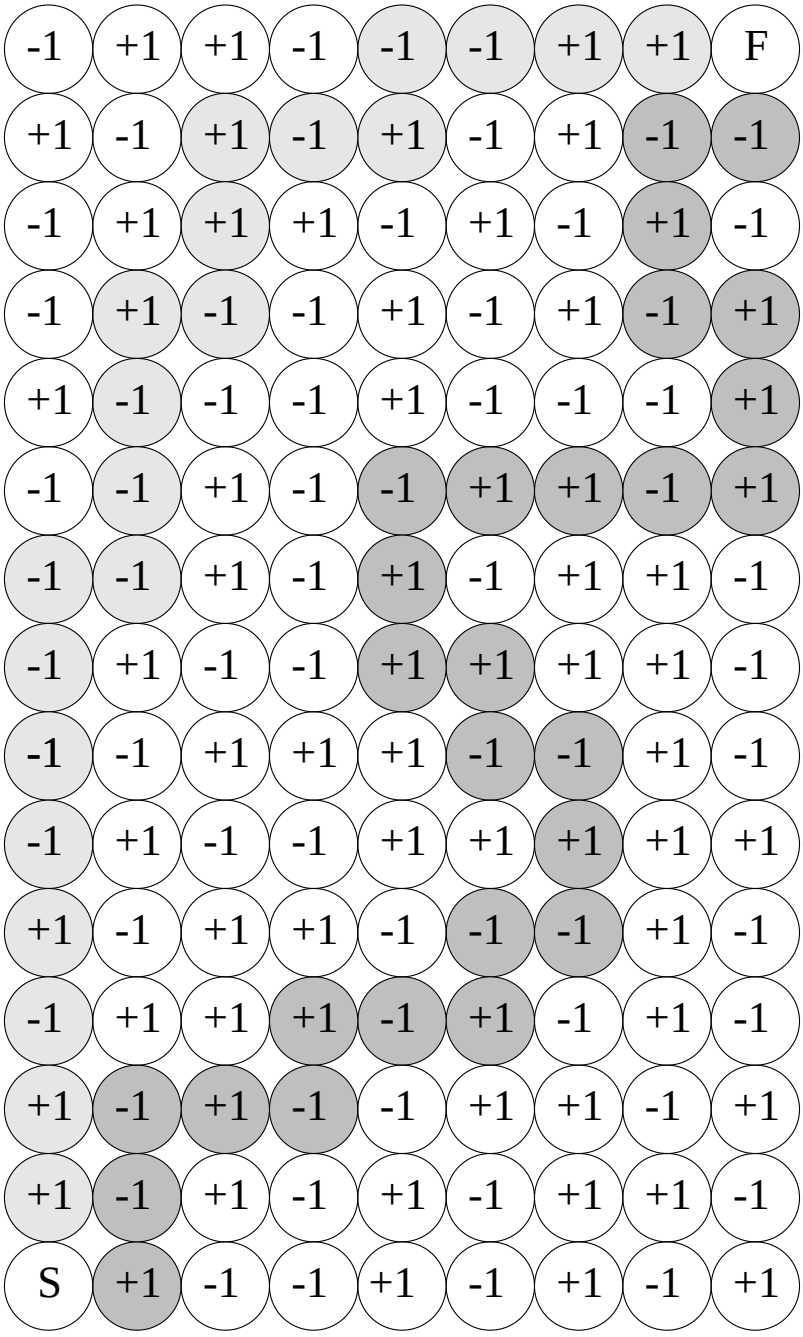

Figure 1.1: Fair Coin

Look at the darker path. If we follow this path and add up the +1s and -1s we find the sums

$$+1, 0, -1, 0, -1, 0, -1, 0, -1, -2, -1, -2, -3, -2, -1, 0,$$
$$-1, 0, +1, 0, +1, +2, +3, +2, +3, +2, +1.$$

All of these sums lie in the shorter interval $[-3, 3]$.

Both of these paths are *self-avoiding paths*. This means that they never visit the same circle twice. Paths may only go from one circle to another if the circles touch.

Puzzle 1: Can you find a self-avoiding path with sums in the interval $[-1, 1]$? Is there a self-avoiding path with sums in a shorter interval?

Next look at Figure 1.2. You see the same grid of circles, but now there are more +1s and less -1s. To make this picture I used a biased coin, one which came up heads roughly six times in seven. We can again add up the +1s and −1s along the light grey path to give

$$+1, +2, +1, +2, +3, +2, +3, +4, +5, +4, +3,$$
$$+4, +3, +4, +5, +4, +5, +6, +7, +8, +9.$$

We see that all of the sums lie in the interval $[0, +9]$. All the sums along the dark grey path lie in the interval $[0, +17]$.

Puzzle 2: Which self-avoiding path from from (S) to (F) in Figure 1.2 has sums in the shortest interval?

We see that it is harder to keep our sums in a short interval if the coin is heavily biased in favour of heads. We are forced to have higher and higher sums. There just aren't enough −1s to cancel out all the +1s we have to meet.

In the first article in this thesis, we explore a larger version of this problem. We cover an infinitely big table with circles, in the same pattern as before. Again we flip a coin and put a +1 or −1 in each circle. We choose a circle, call it the origin, and look at *infinitely* long self-avoiding paths starting from there.

Again we add up the +1s and −1s as we go along these paths. We want to know whether the probability that we can find an infinitely long self-avoiding path, with all sums in a certain interval, could be more than zero. It turns out that:

- Whatever the chance that we see a head, we can never have an infinite self-avoiding path with all its sums in the interval $[0, 1]$.

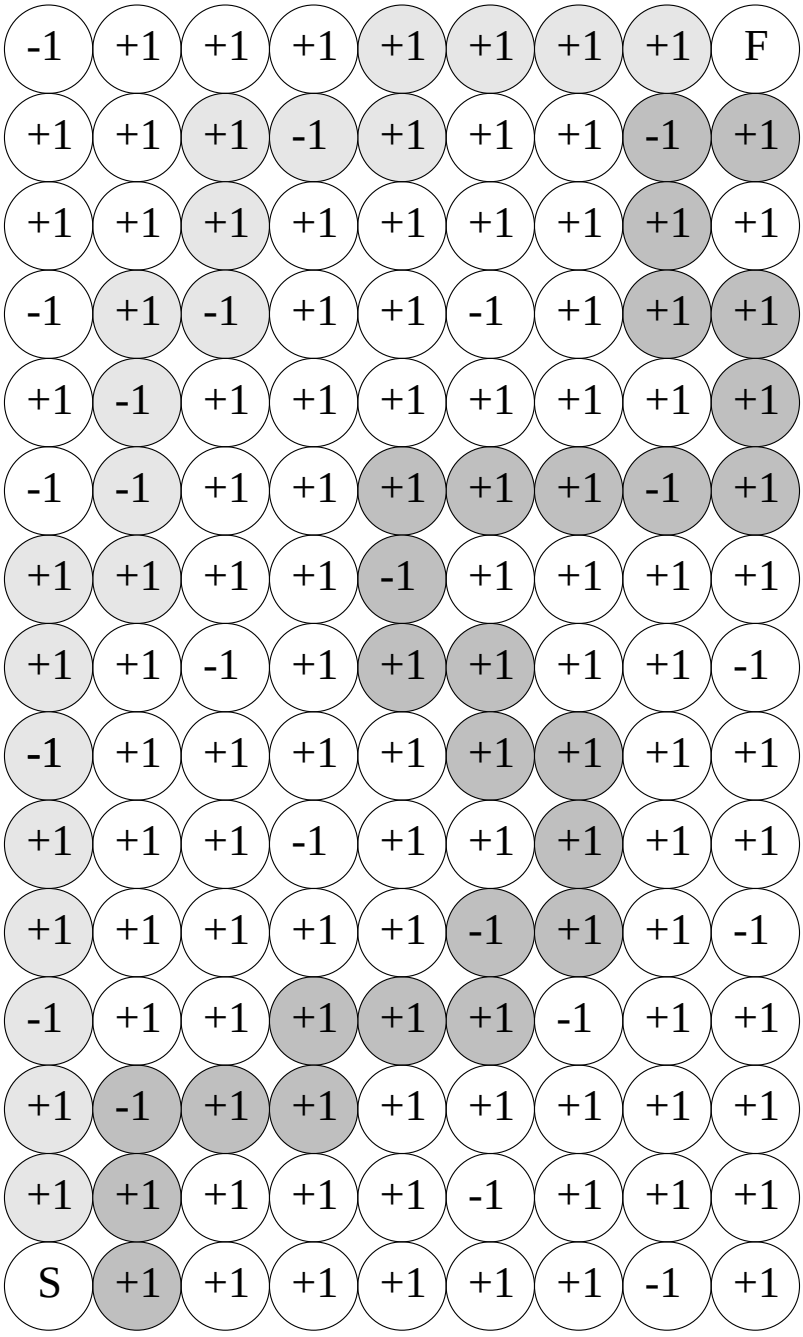

Figure 1.2: Biased Coin

- If the coin is fair or only has a very slight bias, the chance that we see an infinite self-avoiding path with all its sums in $[-21, 21]$ (for example) is positive.

- If the chance that the coin lands head side up is very high, then there is no infinite self-avoiding path with sums in any finite interval, not even $[-10000000000000000, 1000000000000000]$.

We also explore some variants of this problem. Firstly we can change the type of path we allow. One other type of path is the *directed path*. Here we are only allowed to walk upwards and to the left, through the circles. It turns out that we see similar behaviour for these paths as for self-avoiding paths.

Another type of path are those we call the *just-visiting* paths. These paths are allowed to visit a circle more than once, but may only visit each circle finitely many times. These do not include the path that jumps back and forth forever between the origin and the circle at the left of the origin, for example, but could include any path that jumps between these circles thirty times, say, if it then leaves the area forever. Sums along these paths behave differently to those along self-avoiding paths. As long as the probability that we throw a head is not 0 or 1, there is always an interval that our sums can stay in with positive probability. This was not true for self-avoiding paths when the probability of a head is close to 1.

We also ask whether there exists a self-avoiding path whose sums are 0 infinitely often. If we call the probability we see a head p, then we can show that there exists a special number, p_c, such that

- p_c is bigger than 0 and smaller than $1/2$,

- If p is smaller than p_c or bigger than $1 - p_c$ then there is no infinite self-avoiding path with sums that are zero infinitely often,

- If p is between p_c and $1 - p_c$ then there is an infinite self-avoiding path with sums that are zero infinitely often.

1.2 Telephone problems

Imagine that you are an engineer in a telephone company. You have been given the task of deciding where to put radio masts. Each radio mast can send information to, and receive information from, houses within a certain distance - we say that it *communicates* with them.

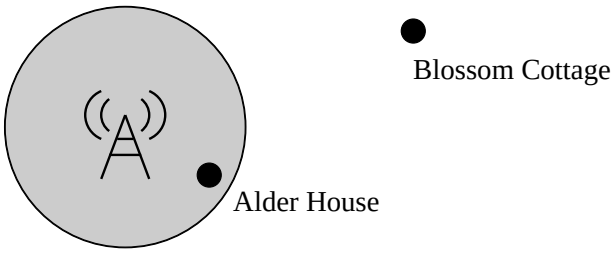

Figure 1.3: a radio mast and two houses

You are given a map of an area, like that in Figure 1.3, Figure 1.6 or Figure 1.7, on which houses are marked as points. In Figure 1.3 you can see a radio mast. The grey disc around it shows the area that it can communicate with. We call this area the *covered area*. For example Alder House can communicate with the mast, as it is inside this covered area, but Blossom Cottage cannot. From now on we will not draw the radio masts but only the grey discs (covered areas). Whenever you see a grey disc you should imagine a radio mast at the centre.

Unfortunately you can't put radio masts wherever you like. Your boss has given you two very important rules to follow:

The Rules

- Every house should be able to communicate with some radio mast. In other words every point should be inside some grey disc.
- Every radio mast should be able to communicate with some house. In other words every grey disc should contain at least one point.

Your boss explains that the first rule ensures all of the customers who live in the house are able to communicate, and so are made happy. Rule two means that the telephone company does not build any telephone mast which will not be used. This saves material and is more environmentally friendly.

Look at Figure 1.4. Here you can see some houses (points), and some covered areas (grey discs). Note that both rules are satisfied.

Problem 1: Can you use less grey discs to cover the points in Figure 1.4? How few can you use? How many do you have to use in Figure 1.7?

There is one thing I have not yet mentioned. This is that radio masts can communicate with each other if their covered areas (grey discs) overlap. This allows houses in different grey discs to send messages to each other. For example, in Figure 1.5 Columbine Cottage can send a message to Dogwood

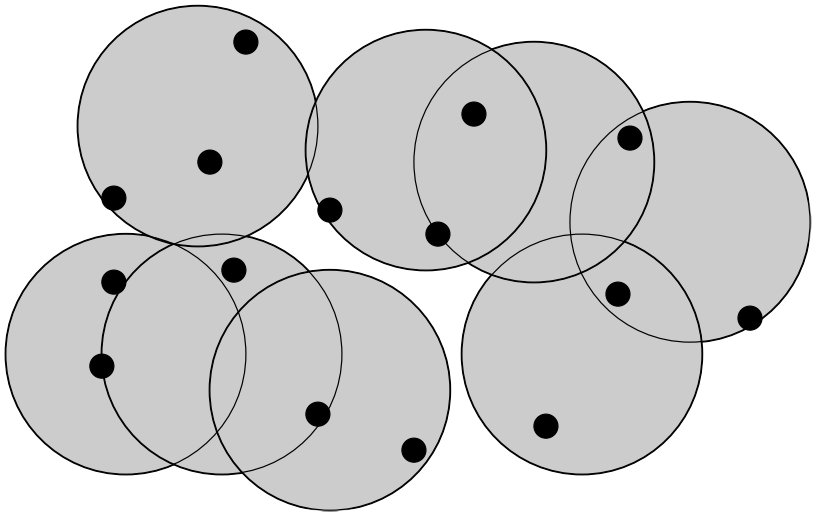

Figure 1.4: Houses and Covered Areas

House, by first sending the message to radio mast one, which sends it to mast three, which sends it to mast two, which sends it to Dogwood House.

Problem 2: Look at Figure 1.6. Can you cover the points according to the rules, so that Elm Cottage can send messages to Feverfew House? How few discs can you use? Is it possible to cover the points in Figure 1.7, so that Gorse Cottage can send messages to Hazel House?

Now I want you to suppose that you dislike the inhabitants of Elm Cottage and Feverfew House. You still have to place radio masts according to the rules, but you are going to try to your very best to stop them being able to send messages to each other.

Problem 3: Can you cover all the points in Figure 1.6, according to the rules, so that Elm Cottage cannot send messages to Feverfew House? What about if you're only allowed to use four radio masts?

In the second article we assume the houses have random positions on an infinite earth, and consider the same kinds of problems. Specifically we assume that the houses are sited according to a certain random process, called a Poisson process.

We call an infinite set of houses that can all communicate with each other an *infinite cluster*. We show that if the density of houses is very small then there can be no infinite cluster. If the density is very high we do not have to make an infinite cluster, unless we have to follow certain extra rules about

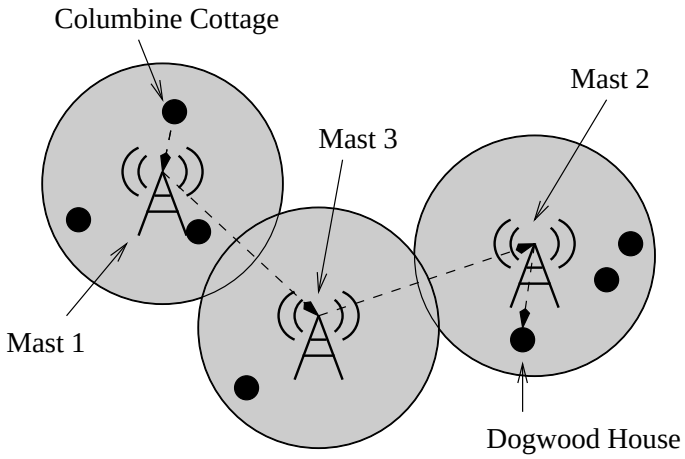

Figure 1.5: a message is sent

where we may place the radio masts. For example, if we may only place the masts on a grid then we will have to make an infinite cluster if the density of houses is large enough. This will also happen if we are only allowed to use a maximum of m masts in each square of area 1, where m is some number.

We also explain the most economical way to place the radio masts. By this I mean that there is no other way that uses a smaller density of masts.

1.3 Infinite trouble

Scientists often want to model big systems. By big I mean that the total size of the system is a lot larger than the things that make up the system. You could think of water molecules in the ocean, or sand grains in a lump of rock.

If water is slowly absorbed into the grass on a sand dune on the edge of Ireland, the movement of water molecules in the centre of the Atlantic is barely affected. Someone kicking Everest doesn't significantly influence the chemical reactions in the middle of the mountain. In this type of situation scientists often choose to pretend that the systems are infinite, in order to avoid questions about what happens on the boundary (the interaction of the edge of the mountain with the man kicking it, the absorption of water into the grass).

One possible problem with this approach is that a model which can be perfectly described when finite, fails to make sense when infinite. This is not always obvious from an informal description of it. In this section I will give three descriptions of systems, and then try to make clear why they don't

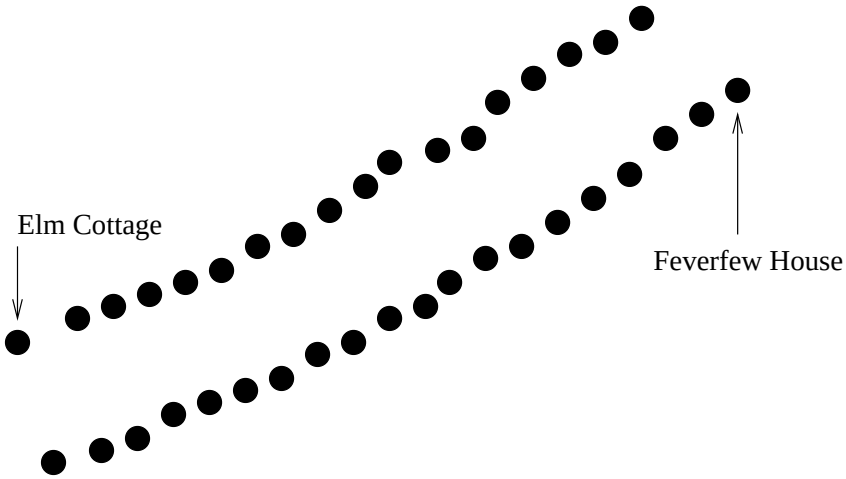

Figure 1.6: a street

Figure 1.7: countryside

work when the system is infinite.

An Example

> In the high street there is a jeweller's shop. The owner of the shop is paranoid about the possibility that someone might steal his diamonds, and so he allows only one customer in his shop at a time. The shop is also only open for one hour a day. Any customer in the shop becomes hypnotised by the diamonds, and will stay until the shop closes.
>
> One Saturday there are k people who want to visit the shop. Each has other errands in the town and so arrives at the shop at a random time in the hour that the shop is open. The time of arrival of one customer doesn't influence the others. If someone arrives at the shop, but sees that there is already a customer inside, they go away disappointed.

This seems like a fine, if strange, story. Notice that each of the k customers has the same chance of being the first to arrive. All of the probabilities must add up to one, so each customer has chance $1/k$ of being the first to arrive. Now suppose that you are one of these customers. The chance that you are the first is also $1/k$. So far so good.

Now let's let k become infinity. This means that infinitely many customers want to go to the shop. If k is infinity the probability that you can go into the shop becomes $1/\infty = 0$. No diamonds today. This is also true for any of the other customers. So who's in the shop?

Another Example

> It's the university sports day, and k of the students who have turned up want to join in the three-legged race. For the three-legged race, students find partners, tie one of each of their legs together and then the pairs race for a hundred metres. Many of them will fall over, hilariously.
>
> However one student, X, is very slow, and nobody wants to tie themselves to X if they have another choice of partner. Can X join in the race?

If k is odd, then all of the other students can find a partner, without including X. If k is even one of them must tie themselves to X. If k is infinity, what happens to X?

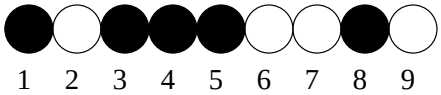

Figure 1.8: A row of coins

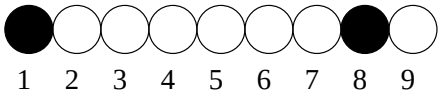

Figure 1.9: The row of coins after 2, 3 and 4 have flipped

Yet Another Example

Suppose we have a row of k fair coins. Each of the coins is coloured white on one side and black on the other. To start with we flip each coin, so that we see the black side of each coin with probability 1/2. If we don't see the black side we will see the white side.

Suppose we see the picture in Figure 1.8, where k is nine. Coin 3 is in a block of three black coins. These three black coins will all turn white, to give the situation in Figure 1.9, after a random time with average $(1/2)^3 = 1/8$ seconds, if none of the coins to either side change colour in this time. In general if there is a block of n coins, all with the same colour, they all change colour after a random time with average length $(1/2)^n$ seconds, if none of the coins to either side change colour in this time. In the situation in Figure 1.9, coin 1 will change colour after an time of average $(1/2)^1$ seconds, as will coins 8 and 9, and coins 2, 3, 4, 5, 6 and 7 will all change colour together after an average time of $(1/2)^6$ seconds, if none of the neighbours of the respective blocks have changed.

Notice that the size of a block will increase if it or one of its neighbouring blocks changes colour, and this will mean that it changes colour faster thereafter.

The problem with this example is that blocks grow very quickly. It is possible to show that the average time until a block of length n has grown to length k (the *total* number of coins) is at most $2 \times (1/2)^n$ seconds. This means that after a random time, of average length at most $2 \times (1/2)^n$, all the blocks become the same colour. Then after an average time of $(1/2)^k$ they

change colour, and they do this again and again.

If k is finite this is all fine. However if k is infinite, something goes wrong. When we start we will have blocks of all lengths, but this implies that the average time until everything becomes one colour is zero. In other words, at the moment we begin, every coins flips to show the same colour. This is different behaviour to that described above above.

These examples may seem strange. However it can be difficult to see that there will be a problem in changing from a finite system to an infinite one when the description is considerably more complicated. Similarly it can be hard to show that there is no problem. Fortunately in the third article we give some conditions under which everything can be made to work.

1.4 Answers

Sum Puzzles

Puzzle 1: There is a self-avoiding path with sums in the interval $[-1, 1]$, but no self-avoiding path with sums in either of the two possible shorter intervals $[-1, 0]$ and $[0, 1]$.

Puzzle 2: I can find a self-avoiding path that stays in $[0, 6]$ (a slight variant of the light grey path), but there might be a better one.

Telephone Problems

Problem 1: You can use less grey discs to cover the points in Figure 1.4 - there is a covering using five discs. In Figure 1.7 there is a covering using two discs.

Problem 2: It is possible to cover the points according to the rules, so that Elm Cottage can send messages to Feverfew House, using four discs. It is, however, not possible to cover the points in Figure 1.7 so that Gorse Cottage can send messages to Hazel House.

Problem 3: It is possible to cover all the points in Figure 1.6, according to the rules, so that Elm Cottage cannot send messages to Feverfew House, but if you must use four radio masts they will be able to send messages to each other. This is one of the interesting phenomena we study in Chapter 3 - by restricting the number of radio masts used, we force communication.

Chapter 2

Infinite paths with bounded or recurrent partial sums

Lorna Booth and Ronald Meester

This chapter has been published in Probability Theory and Related Fields, **120** (2001), no. 1, 118–142.

We consider problems of the following type. Assign independently to each vertex of the square lattice the value $+1$, with probability p, or -1, with probability $1-p$. We ask whether an infinite path π exists, with the property that the partial sums of the ± 1s along π are uniformly bounded, and whether there exists an infinite path π with the property that the partial sums along π are equal to zero infinitely often. The answers to these question depend on the type of path one allows, the value of p and the uniform bound specified. We show that phase transitions occur for these phenomena. Moreover, we make a surprising connection between the problem of finding a path to infinity (not necessarily self-avoiding, but visiting each vertex at most finitely many times) with a given bound on the partial sums, and the classical Boolean model with squares around the points of a Poisson process in the plane. For the recurrence problem, we also show that the probability of finding such a path is monotone in p, for $p \geq \frac{1}{2}$.

2.1 Introduction

In this paper we consider a connected, infinite, locally-finite graph, G, with vertex set V, and edge set E. This will generally be $\mathbb{Z}^2$, although many of our results can be extended to other graphs. One particular vertex is distinguished and called the origin, 0. To every vertex $v \in V$, we assign a random variable X_v, which takes value 1 with probability p and value -1 otherwise, independently of the values at other vertices.

A path from a vertex z_0 say, is a sequence of vertices $\pi = (z_0, z_1, z_2, \ldots)$ such that $(z_{i-1}, z_i) \in E$ for $i = 1, 2, \ldots$ For such a path we define the partial sums $S_n^\pi = \sum_{i=1}^n X_{z_i}$ for $n = 1, 2, \ldots$, where, for our later convenience, we do not count the value at the starting point, z_0. We are concerned with the question of whether there exists an infinite path π with either $\sup_n\{|S_n^\pi| < \infty\}$ or with $S_n^\pi = 0$, for infinitely many n. Note that these problems are symmetric in p around $\frac{1}{2}$. The answers to these questions will depend upon the types of paths we allow ourselves to use. We will consider three different types of infinite paths:

Definition *On the square lattice, $\mathbb{Z}^2$, an* **oriented path**, *$\pi = (z_0, z_1, \ldots)$, is a path such that $z_i - z_{i-1} = e_1$ or $z_i - z_{i-1} = e_2$, for $i = 1, 2, \ldots$, where e_1 and e_2 are the first and second coordinate vectors. A* **self-avoiding path** *is a path whose vertices are all distinct. Finally, a* **just-visiting path** *is a path in which any vertex appears only finitely many times.*

All oriented paths are self-avoiding, and all self-avoiding paths are just-visiting. Note also that the existence of an infinite self-avoiding path with bounded partial sums starting at some vertex, implies that there exists one from every other vertex. Such a path can be constructed, for example, by taking the shortest self-avoiding path from the chosen vertex to the bounded path and thereafter following it. This also true for just-visiting paths, similarly.

On the integer line (where integers are connected by an edge if their difference is one), the answers to these questions for self-avoiding paths are well known, as they refer to the simple random walk. (For information about this see for example [10].) Here we have that when $p = \frac{1}{2}$ we have no bounded partial sums, but we do have partial sums that are zero infinitely often, almost surely. At all other values of p we have neither behaviour. If we consider just-visiting paths, it can be shown that there are no paths with bounded partial sums, almost surely, for any value of p, but for all $p \in (0, 1)$ there are just-visiting paths with partial sums that are zero infinitely often.

Benjamini and Peres [6], answered both questions (and many other more general ones), for self-avoiding paths on trees. Given a tree, T, the boundary,

∂T is the set of rays or infinite self-avoiding paths emanating from the root, 0. If we denote by $\dim(\partial T)$ the Hausdorff dimension of this boundary (see [8] for an explanation of Hausdorff dimension) then their theorem states that a path from 0 with bounded partial sums exists with positive probability if and only if $\dim(\partial T) > \log(\frac{1}{2\sqrt{p(1-p)}})$. If we denote the packing dimension by $\mathrm{Pdim}(\partial T)$ (see [8]), then they show that for $\mathrm{Pdim}(\partial T) < \log(\frac{1}{2\sqrt{p(1-p)}})$ there are no self-avoiding paths with partial sums returning to zero infinitely often, almost surely, while for $\dim(\partial T) > \log(\frac{1}{2\sqrt{p(1-p)}})$ these exist with positive probability.

A sufficient condition to have infinite paths with both bounded partial sums and partial sums that are zero infinitely often, is the existence of AB percolation, for appropriate parameter values. From our viewpoint AB percolation asks for the almost sure existence of a self-avoiding path with alternate 1s and -1s. Thus the partial sums are zero at every other point along the path, and never exit either of the intervals $[0, 1]$ or $[-1, 0]$. It has been shown (for example) that AB percolation occurs on the triangular lattice for an interval of values of p around $1/2$, see [2]. AB percolation has been shown not to occur, for any value of p, on the square and hexagonal lattices, see [3]. We note that AB percolation occurs for self-avoiding paths if and only if it occurs for just-visiting paths. An interesting feature of AB percolation is that its probability is not monotonic in p on $[1/2, 1]$ for many graphs, see [15].

A related, yet weaker, question is whether ρ-percolation occurs, that is, whether there exists a self-avoiding path π such that $\liminf_{n\to\infty} \frac{S_n^\pi}{n} \geq 2\rho - 1$ (see [17]). AB percolation, standard percolation, ρ-percolation and the paths we investigate are all special cases of the general question of which words (infinite sequences of -1s and 1s) can be seen along self-avoiding paths in a percolation configuration (see [5], [12] and [13]).

We end this section with some notation and definitions used throughout. The product measure described above is denoted by $\mathbb{P}_p$. Denote by $\mathbb{E}_p$ the corresponding expectation operator. We call two vertices *adjacent* if there is an edge between them, and we call two edges adjacent if they share a vertex. We define the distance between two points, $v^1 = (v_1^1, v_2^1), v^2 = (v_1^2, v_2^2) \in \mathbb{R}^2$ as $|v_1^1 - v_1^2| + |v_2^1 - v_2^2|$, the L_1-distance. An interval in $\mathbb{R}$ is said to have *size* s if it contains exactly s integers.

The next section in this paper contains our principal results, and the subsequent sections contain the proofs.

2.2 Principal results

Our first theorem states that there is a non-trivial phase transition for oriented paths.

Theorem 2.2.1 *On the square lattice the following hold.*
(a) There are no infinite oriented paths with sums bounded in an interval of size $M \geq 2$, $\mathbb{P}_p$-almost surely, for p in the set:

$$\left[0, \frac{1-\sqrt{1-1/4\cos^2(\pi/(M+1))}}{2}\right) \cup \left(\frac{1+\sqrt{1-1/4\cos^2(\pi/(M+1))}}{2}, 1\right].$$

In particular, for $p \in \left[0, \frac{1}{2}-\frac{1}{2}\sqrt{\frac{3}{4}}\right) \cup \left(\frac{1}{2}+\frac{1}{2}\sqrt{\frac{3}{4}}, 1\right]$, there are no paths with sums bounded in any interval, $\mathbb{P}_p$-almost surely.
(b) There are, $\mathbb{P}_p$-almost surely, infinite oriented paths with partial sums that return to zero every 42 steps for $p \in (0.475, 0.525)$. This immediately implies that there are oriented paths with partial sums that are bounded and that return to zero infinitely often, $\mathbb{P}_p$-almost surely, and thus also self-avoiding and just-visiting paths with the same properties.

We see here a contrast with AB percolation, which occurs for no values of p on this lattice.

We need a separate statement to rule out the possibility of infinite self-avoiding paths with bounded partial sums, or with partial sums that return to 0 infinitely often, for p close to 0 and 1. We shall formulate the next result for more general graphs.

For a graph G, let $\sigma_G(n)$ be the number of self avoiding walks of length n, from the origin. Let $c_G = \lim_{n\to\infty} \sigma_G(n)^{1/n}$, if this exists. c_G is called the connectivity constant of G. For the integer lattice (and many others), existence of the connectivity constant follows from subadditivity.

Theorem 2.2.2 *Let G be a graph with a well-defined connectivity constant c_G. Then there are, $\mathbb{P}_p$-almost surely, no infinite self-avoiding paths with partial sums that return to 0 infinitely often or with partial sums that are bounded in some interval for*

$$p \in \left[0, \frac{1-\sqrt{1-\frac{1}{c_G^2}}}{2}\right) \cup \left(\frac{1+\sqrt{1-\frac{1}{c_G^2}}}{2}, 1\right].$$

For the square lattice it has been shown rigorously that $c_{\mathbb{Z}^2} \leq 2.7$ (see [1]), which implies that we have no infinite self-avoiding path with bounded partial sums for $p \in [0, 0.035) \cup (0.965, 1]$.

For just-visiting paths, the situation is quite different. For every $p \in (0,1)$ we can define a minimal interval size, $I(p)$ say, to be the minimal integer, such that there exists a just-visiting path from the origin, with partial sums bounded in an interval of length $I(p)$, with positive probability.

Theorem 2.2.3 *On the square lattice the following hold.*
(a) For any $p \in (0,1)$ we have $2 \leq I(p) < \infty$. That is, for any $p \in (0,1)$ there is, with positive $\mathbb{P}_p$ probability (and hence with $\mathbb{P}_p$ probability one), an infinite just-visiting path with bounded partial sums. We can take this path so that the partial sums are equal to 0 infinitely often.
(b) We have

$$\lim_{p \to 1} I(p) = \infty.$$

That is, for any interval J on the real line we can find p, close to 1, such that for this value of p, no just-visiting paths with all partial sums in J exists, $\mathbb{P}_p$-almost surely.

Our next result gives a connection between two apparently unrelated processes, the percolation of just-visiting paths with bounded sums, and the classical Boolean model.

Consider a Poisson process of rate λ in the plane and centre a diamond (with fixed orientation) of radius $1/2$ at every Poisson point, that is, v is in the diamond centred at a Poisson point x if the distance between v and x is at most $1/2$. This process is known to have a critical point for percolation λ_c, such that for $\lambda \leq \lambda_c$ there is no percolation (that is, the union of all diamonds contains no unbounded connected component almost surely), while for $\lambda > \lambda_c$ percolation occurs. Later we shall need to talk about diamond processes formed by placing a diamond of radius r at each of the points of the Poisson process, and we denote the critical point for this model by $\lambda_c(r)$. It can be seen by a simple scaling argument that,

$$r_1^2 \lambda_c(r_1) = r_2^2 \lambda_c(r_2) \tag{2.1}$$

for any two values, r_1 and r_2. See [16] for information on this type of model, and the results mentioned above.

Theorem 2.2.4 *Consider a sequence $q_1, q_2, q_3, \ldots$ of numbers in $(0,1)$ converging to 1.*

(a) If $\limsup_{N\to\infty} 2N^2(1-q_N)^2 < \lambda_c$ *then, for* N *sufficiently large, there* $\mathbb{P}_{q_N}$*-almost surely does not exist an infinite just-visiting path with partial sums bounded in* $[0, N]$.
(b) If either

$$\lambda_c < \liminf_{N\to\infty} 2N^2(1-q_N)^2 \le \limsup_{N\to\infty} 2N^2(1-q_N)^2 < \infty$$

or

$$\liminf_{N\to\infty} 2N^2(1-q_N)^2 = \infty,$$

then for N *sufficiently large there exists, with positive* $\mathbb{P}_{q_N}$ *probability (and hence with* $\mathbb{P}_p$ *probability one), an infinite just-visiting path with partial sums bounded in* $[0, N]$.

Two remarks are appropriate here. In the first place, we could replace the interval $[0, N]$ in this theorem by any sequence of intervals $J_1, J_2, \ldots$, with J_k of size k. We chose the above formulation for the sake of simplicity of the statements. Secondly, the results in (a) and (b) together will imply that we can strengthen the result of Theorem 2.2.3(b) in the following sense:

Corollary 2.2.5 *The interval size* $I(p)$ *defined before Theorem 2.2.3 satisfies*

$$\lim_{p\to 1}(1-p)I(p) \to \sqrt{\frac{\lambda_c}{2}}.$$

Our next result gives a relationship between paths with partial sums that return to zero infinitely often and those with partial sums that do not converge to $+\infty$. It is motivated by the question of whether the probability of an infinite self-avoiding path with bounded partial sums is monotone in $p > 1/2$. This we do not know, but we can prove a monotonicity statement for the probability of having an infinite self-avoiding path with partial sums equal to zero infinitely often. This quickly follows from the next result.

Theorem 2.2.6 *On the square lattice, for any value of* $p > 1/2$ *such that with positive probability, there is an infinite self-avoiding path with partial sums that do not converge to* $+\infty$*, there is, with positive probability, an infinite self-avoiding path with partial sums that return to zero infinitely often.*

Corollary 2.2.7 *On the square lattice, for* $p > 1/2$*, the probability that there exists an infinite self-avoiding path with partial sums that are equal to zero infinitely often is monotone decreasing in* p.

2.3 Bounded sums on self-avoiding paths

The first part of Theorem 2.2.1 is proved via a recurrence method, the second part with a second moment method.

Proof of Theorem 2.2.1(a) We will show that the expected number of oriented paths of length n, from any point on the line $x + y = 0$ to a given point (a, b) (with $a + b = n,\ a, b \in \mathbb{Z}$), with sums bounded in the interval, tends to zero as n tends to infinity. To see that this suffices, note that for any interval I, the expected number of paths from the origin of length n with partial sums in I is equal to the expected number of paths to the point $(a, n - a)$ from the line $x + y = 0$ with partial sums in I, for any $a \in \mathbb{Z}$.

Given an interval, I, of size M, and a point (a, b) (with $a + b > 0$) define $V_{(a,b)}$ to be the random vector with elements $v_{(a,b),i}, i \in I$ which record the number of oriented paths from $x + y = 0$ to (a, b) which have partial sums bounded in the interval and final partial sum i.

Let $\Lambda_{(a,b)}$ be the random $M \times M$ matrix that has entries,

$$\lambda_{ij} = \begin{cases} 1 & i = j - 1, \\ 0 & \text{otherwise}, \end{cases}$$

if (a, b) is assigned value $+1$, and

$$\lambda_{ij} = \begin{cases} 1 & i = j + 1, \\ 0 & \text{otherwise}, \end{cases}$$

if (a, b) is assigned value -1.

Then,

$$V_{(a,b)} = \Lambda_{(a,b)}(V_{(a-1,b)} + V_{(a,b-1)}).$$

We now take the expectation of both sides of this equation, noting two things. The first is that $\Lambda_{(a,b)}$ is independent of $V_{(a-1,b)}$ and $V_{(a,b-1)}$. The second is that the expectation of $V_{(a,b)}$ is the same as that of $V_{(c,d)}$ if $a + b = c + d$. Thus we may write $\bar{V}_n$ to be the expectation of any $V_{(a,b)}$ with $a + b = n$. Let $\bar{\Lambda}$ be the expectation of $\Lambda_{(0,0)}$ (which is also the expectation of $\Lambda_{(a,b)}$ for any (a, b), as they are identically distributed). Hence we have that

$$\bar{V}_n = \bar{\Lambda}(\bar{V}_{n-1} + \bar{V}_{n-1}) = 2^n \bar{\Lambda}^n \bar{V}_0.$$

From this we see that if the largest absolute value of an eigenvalue of $2\bar{\Lambda}$ is less than 1, $\lim_{n\to\infty} \bar{V}_n$ will be the zero vector. This then implies that with probability one the number of infinite paths with sums bounded in the

interval is zero. Now if we denote the entries of $\bar{\Lambda}$ by $\bar{\lambda}_{ij}$, we have that,

$$\bar{\lambda}_{ij} = \begin{cases} p & i = j-1, \\ 1-p & i = j+1, \\ 0 & \text{otherwise.} \end{cases}$$

It can be easily calculated that the largest eigenvalue of this matrix is

$$2\sqrt{p(1-p)}\cos(\pi/(M+1)),$$

which gives the result. □

For the proof of Theorem 2.2.1(b) we need the following combinatorial lemma, which we will prove at the end of this section. We denote by N_n the number of oriented paths from $(0,0)$ to $(n,2n)$ with final partial sum zero.

Lemma 2.3.1 *We have, for n even,*

$$\mathbb{P}_p(N_n > 0) \geq \frac{\binom{3n}{n}^2 \binom{3n}{3n/2}^2}{\sum_{k=1}^{3n} T(n,k) \sum_{i=0}^{k} \binom{k}{i} \binom{3n-k}{3n/2-i}^2 p^{-i}(1-p)^{-k+i}},$$

where

$$T(n,k) = \sum_{a=0}^{3n/2} \sum_{m=0}^{k} \binom{3n-2a}{n-a} 2^{2a} C(a,m) \binom{k}{m} \binom{3n-k-1}{3n-2a-(k-m)},$$

and $C(a,m)$ is the coefficient of t^{2a} in the series expansion of $(1-\sqrt{1-t^2})^m$. With $n = 14$ and $p \in (0.475, 0.525)$ this gives us a lower bound of 0.934.

Proof of Theorem 2.2.1(b) Lemma 2.3.1 shows that the probability of diagonally crossing a $n \times 2n$ rectangle, with a final partial sum of zero is high. We then combine several of these events in a particular way, with a probability that is still high. Finally we place these rectangles onto a larger grid in such a way that we can show that this stochastically dominates supercritical ordinary site percolation on the square lattice. This then implies the existence, with positive probability, of a path with the required property.

We define, for even n, the event C_1 as the event that there exists an oriented path from $(0,0)$ to $(n,2n)$ with final partial sum 0, and $C_2, \ldots, C_4$ similarly with $(0,0)$ to $(n,2n)$ replaced by $(0,0)$ to $(2n,n)$, $(-2n,n)$ to $(0,0)$ and $(-n,-2n)$ to $(0,0)$ respectively. The event $C = C_1 \cap C_2 \cap C_3 \cap C_4$ then has probability at least $0.753 > 3/4$, for $n = 14$ and $p \in (0.475, 0.525)$ (using the fact that the events C_1 and C_2 are independent of C_3 and C_4). It is

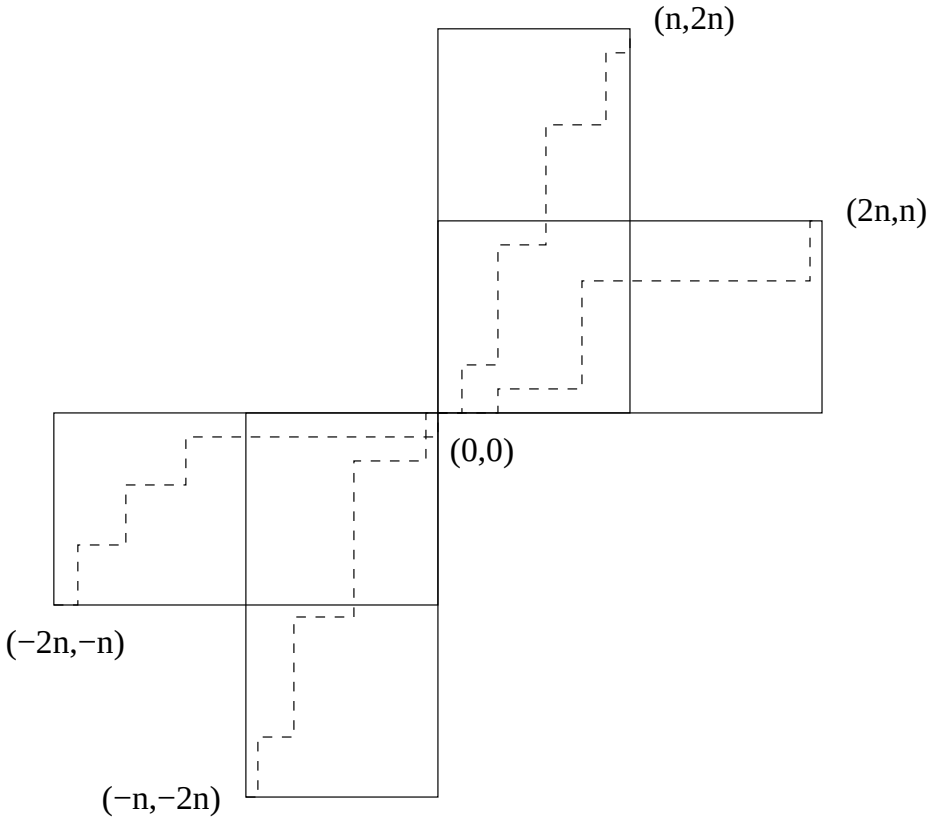

Figure 2.1: The event C, dashed lines indicate paths that cross the rectangles with final partial sums of zero.

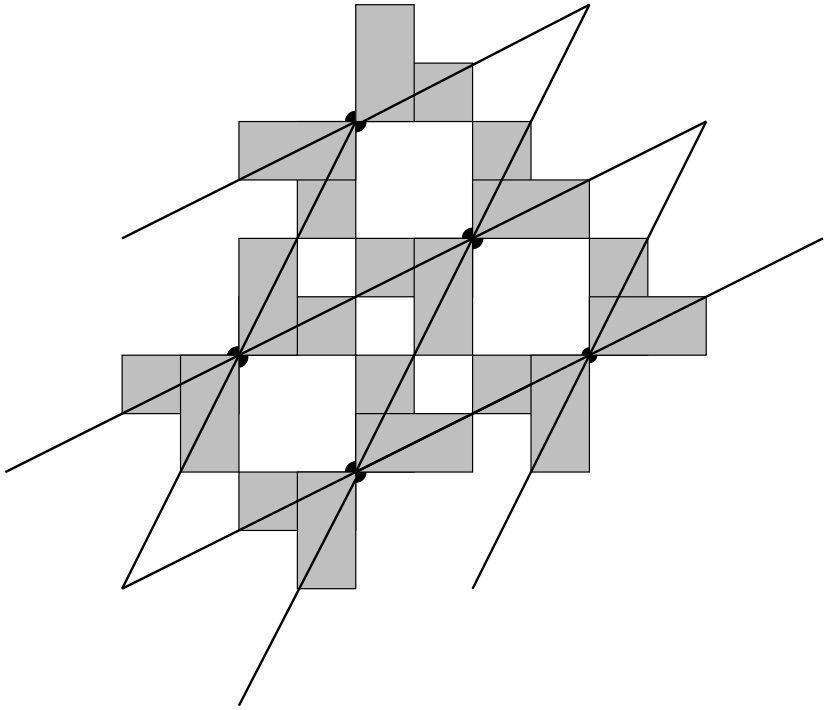

Figure 2.2: The larger grid; the grey areas are those we consider, heavy lines indicate the edges of the larger distorted lattice.

illustrated in Figure 2.1. We denote by $w + C$ the event that C occurs, but translated to w from $(0, 0)$.

We now move to the larger grid formed by the vertices $\{(2an, an) + (bn, 2bn) : a, b \in \mathbb{Z}\}$ with edges from $(2an, an) + (bn, 2bn)$ to $(2(a + 1)n, (a + 1)n) + (bn, 2bn)$ and $(2an, an) + ((b + 1)n, 2(b + 1)n)$. Note that this grid is a distorted version of the square lattice, and so has the same critical point for oriented independent percolation, which is known to be no more than 3/4 (see [14]). We call a vertex, $(2an, an) + (bn, 2bn)$, on this grid 'open' if the event $(2an, an) + (bn, 2bn) + C$ occurs - see Figure 2.2.

A path of open vertices on the large grid then implies a path on the standard grid with partial sums that are zero every $3n$ steps. This certainly happens with positive probability for $n = 14$ and $p \in (0.475, 0.525)$. □

Proof of Lemma 2.3.1 We shall estimate the probability that $N_n > 0$ via a second moment method. The expected number of oriented paths from $(0, 0)$ to $(n, 2n)$ with final partial sum zero is

$$\binom{3n}{n} \times \binom{3n}{3n/2} p^{3n/2}(1 - p)^{3n/2}.$$

Given two of these paths, if they meet at precisely k vertices (excluding $(0,0)$), we claim that the probability that they both have partial sum zero at $(n,2n)$ is

$$\sum_{i=0}^{k} \binom{k}{i} \binom{3n-k}{3n/2-i}^2 p^{3n-i}(1-p)^{3n-k+i}. \tag{2.2}$$

To see this, we condition upon the sum along the shared portion of the two paths. The probability this sum is equal to j is $\binom{k}{\frac{k+j}{2}} p^{(j+k)/2}(1-p)^{(k-j)/2}$, when $(k+j)/2$ is an integer, and zero otherwise. The sums along the non-shared portions of the two paths are independent and each has probability

$$\binom{3n-k}{\frac{3n-k-j}{2}} p^{(3n-k-j)/2}(1-p)^{(3n-k+j)/2}$$

of being equal to $-j$. Thus summing over the possible values of j we find that the probability that both paths have partial sum zero at $(n,2n)$ is equal to

$$\sum_{\substack{j=-k,\\ j \text{ with same parity as } k}}^{k} \binom{k}{\frac{k+j}{2}} \binom{3n-k}{\frac{3n-k-j}{2}}^2 p^{(3n-\frac{(k+j)}{2})}(1-p)^{(3n-\frac{(k-j)}{2})}.$$

If we now substitute i for $\frac{j+k}{2}$, we find formula (2.2) above.

Thus to calculate the expected value of N_n^2, we need to know only how many (ordered) pairs of oriented paths from $(0,0)$ to $(n,2n)$ there are that meet in precisely k places (excluding $(0,0)$). If we denote this number by $T(n,k)$ then we have that,

$$\mathbb{E}_p(N_n^2) = \sum_{k=1}^{3n} T(n,k) \sum_{i=0}^{k} \binom{k}{i} \binom{3n-k}{3n/2-i}^2 p^{3n-i}(1-p)^{3n-k+i}.$$

We first show that $T(n,k)$ is equal to the number of paths in $\mathbb{Z}^2$ with a certain property. Consider a pair (π,π') of oriented paths in $\mathbb{Z}^2$ from $(0,0)$ to $(n,2n)$. We map (call this map A) this pair to an undirected path $(z_0,z_1,\ldots,z_{3n})$ in $\mathbb{Z}^2$ as follows. Start in the origin, that is, $z_0=(0,0)$. The path is now constructed sequentially as follows. If π and π' both make a step to the right, then $z_1 = z_0 + e_1$; if π and π' both make a step upward, then $z_1 = z_0 - e_1$; if π goes up, and π' goes to the right, then $z_1 = z_0 + e_2$; if π goes to the right and π' goes upwards, then $z_0 = z_1 - e_2$. This procedure is repeated for each of the steps of π and π'. Here, e_1 and e_2 denote the unit vectors. For example, if $\pi = ((0,0),(1,0),(1,1),(1,2))$ and $\pi' = ((0,0),(0,1),(0,2),(1,2))$ then we find the two dimensional path shown in Figure 2.3. Let us define

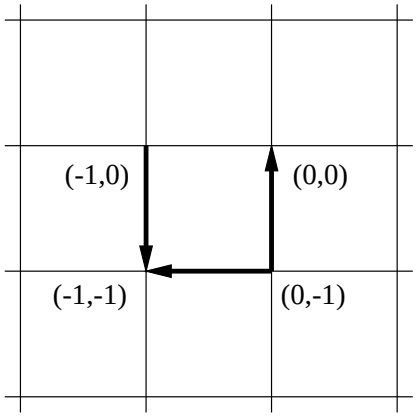

Figure 2.3: Path produced from $\pi = ((0,0),(1,0),(1,1),(1,2))$ and $\pi' = ((0,0),(0,1),(0,2),(1,2)$ by map A or from $x = (0,-1)$, $y = (0,-1,0)$, and $S = (v,h,v)$ by map B.

a *relevant path of length* $3n$ in $\mathbb{Z}^2$ to be a series of vertices $z_0, z_1, z_2, \ldots, z_{3n}$ with $z_0 = (0,0)$, $z_{3n} = (-n,0)$ and $\|z_i - z_{i-1}\| = 1$, $i = 1,2,\ldots,3n$. Define a k-path of length $3n$ as a relevant path of length $3n$ in which precisely k of the z_i, $i = 1,2,\ldots,3n$, have second coordinate 0. The map A described above maps a pair of paths (π,π') that meet in k places to a k-path in a bijective fashion. We conclude that $T(n,k)$ is the number of k-paths of length $3n$, and we shall now explain how we can find this number.

Given two one dimensional sequences, $y = (y_0, y_1, \ldots, y_{2a})$ (of length $2a$) and $x = (x_0, x_1, \ldots, x_{3n-2a})$ (of length $3n-2a$), both starting in 0, and making steps of size ± 1, and a sequence $S = (s_1, s_2, \ldots, s_{3n})$, $s_i \in \{v,h\}$, $i = 1,2,\ldots,3n$ containing $2a$ v's and $3n-2a$ h's we can construct a path $z = (z_0, z_1, \ldots, z_{3n})$ of length $3n$, in the following way. Start in the origin. We copy steps in x to horizontal steps in z and steps in y to vertical steps in z, and the order in which this is done is governed by the order of the letters in S. For instance, if we take $x = (0,-1)$, $y = (0,-1,0)$, and $S = (v,h,v)$ we again find the path in Figure 2.3. We call this map B. It is easy to see that every relevant path can be produced from this procedure given the appropriate choice of x, y and S, and that B is injective.

Map B produces a k-path of length $3n$ from $y = (y_0, y_1, \ldots, y_{2a})$, $x = (x_0, x_1, \ldots, x_{3n-2a})$ and S if and only if all of the following hold, for some $m \leq k$:

1. $y_0 = y_{2a} = 0$ and m of the steps of y end at zero, that is, the set $\{0 < i \leq 3n-2a\,;\, y_i = 0\}$ contains m elements;

2. $x_0 = 0$, $x_{3n-2a} = -n$;

3. $k - m$ of the steps of y either occur after a step of x that ended at zero but before another step of x, or before any steps of x.

We shall now count how in how many ways this is possible. The probability generating function for the first return time for a one dimensional simple symmetric random walk is given by $(1 - \sqrt{1-t^2})$. Thus the probability that such a random walk returns to the origin for the mth time at time $2a$ is the coefficient of t^{2a} in $(1-\sqrt{1-t^2})^m$. Denote this coefficient by $C(a,m)$. There are 2^{2a} such simple symmetric random walk paths of length $2a$, and they are all equally likely. Hence the number of paths y above of length $2a$ with precisely m steps that end at zero is is $2^{2a}C(a,m)$.

There are $\binom{3n-2a}{n-a}$ choices for x. Given x and y, we need to find how many sequences S of h's and v's give k-paths. Now S must contain $2a$ v's, and we identify these, in order, with the $2a$ steps of y and label those that correspond to steps in y that end at 0 with a star. We need to fit the remaining $3n-2a$ h's into this sequence, such that $k-m$ of them appear either before any v, or after a starred v but before any other v, and the other $3n-2a-(k-m)$ appear in other places. We claim that we can do this in $\binom{k}{m}\binom{3n-k-1}{3n-2a-(k-m)}$ ways. This is the product of the number of ways in which $(k-m)$ h's can be fitted into the $(m+1)$ places that are either before any v, or after a starred v but before any other v (denote this number by $W(k-m, m+1)$) and the number of ways in which the remaining h's can be fitted into the remaining $2a-m$ places. To calculate $W(k-m, m+1)$ we work out the number of ways in which the $(k-m)$ h's could be fitted into the $(m+1)$ places if the h's were labelled $h_1, h_2, \ldots, h_{k-m}$, and then divide by the number of labellings, $(k-m)!$ We can put h_1 into $(m+1)$ places. There are then $(m+2)$ places into which we can put h_2, directly to either side of h_1 and into the places in which we did not put h_1. There are, similarly, $(m+3)$ places where we can put h_3, and as we continue in this manner, there will be $(m+i)$ places in which to put h_i. Hence

$$W(k-m, m+1) = \frac{(m+1)(m+2)\ldots(m+(k-m))}{(k-m)!} = \binom{k}{m}.$$

Counting the number of ways in which we can fit the rest of the h's into the other places works in exactly the same way. Thus if we sum over m and a we find the total number of k-paths,

$$T(n,k) = \sum_{a=0}^{n}\sum_{m=0}^{k}\binom{3n-2a}{n-a}2^{2a}C(a,m)\binom{k}{m}\binom{3n-k-1}{3n-2a-(k-m)}.$$

Finally, since N_n is non-negative and integer-valued, the second moment bound

$$\mathbb{P}(N_n > 0) \geq \frac{\mathbb{E}_p(N_n)^2}{\mathbb{E}_p(N_n^2)}$$

then gives the desired result. □

2.4 Bounded and recurrent sums

The proofs of Theorems 2 and 3 are not difficult and are based on simple counting arguments. We will, however, need the following lemma. The proof of the lemma comes after the proof of the theorem.

Lemma 2.4.1 *If, for a given value of p, there is an infinite self-avoiding path with bounded partial sums, with positive probability, then there is an infinite self-avoiding path whose partial sums returns to 0 infinitely often, with positive probability.*

Proof of Theorem 2.2.2 We start by noting that if there exists a path with partial sums that return to 0 infinitely often, with positive probability, then the expected number of points that can be reached from 0 by a self-avoiding path with final partial sum 0 must be infinite. However the expected number of points that can be reached from 0 by a self-avoiding path with final partial sum 0 is certainly no greater than the expected number of finite self-avoiding paths that have partial sum 0 on their last step. The number of self-avoiding paths of length n is $\sigma_G(n)$, so this expectation is

$$\begin{aligned} & \sum_{n>0,\ n \text{ even}} \sigma_G(n) \binom{n}{n/2} p^{n/2}(1-p)^{n/2} \\ \leq & \sum_{n>0,\ n \text{ even}} \sigma_G(n) 2^n (p(1-p))^{n/2} < \infty \end{aligned}$$

for p such that $2^2 c_G^2 p(1-p) < 1$. Hence for these values of p (which are those given in the theorem) there can be no path with partial sums that return to 0 infinitely often. The part of the theorem referring to bounded partial sums then follows immediately from Lemma 2.4.1. □

Proof of Lemma 2.4.1 We first show that

$$\mathbb{P}_p(\exists \text{ a path } \pi \text{ with } S_n^\pi = 0 \text{ i.o.}) = 0 \tag{2.3}$$

implies that for all $z \in \mathbb{Z}$,

$$\mathbb{P}_p(\exists \text{ a path } \pi \text{ with } S_n^\pi = z \text{ i.o.}) = 0. \tag{2.4}$$

To see this, suppose that an integer, z say, exists such that the event in (2.4) has positive probability. Then there must exist a non-random vertex, x say, such that with positive probability, there exists a path which has partial sums that are z infinitely often, passes through x, and has partial sum z at x. Then we have a path from x that has partial sums that are zero infinitely often. This contradicts (2.3).

Now, if we have an infinite self-avoiding path with partial sums that are bounded, then these partial sums must visit some value infinitely often. Thus, by the previous paragraph, if we have an infinite self-avoiding path with partial sums that remain bounded, with positive probability, we must have an infinite self-avoiding path with partial sums that are zero infinitely often, with positive probability. □

Proof of Theorem 2.2.3 To prove the first part of this theorem we first fix $p \in (0,1)$. We can find a box size, m, such that with probability higher than 3/4 the box $B(m) = \{(x,y) : x,y \in \{0,1,\ldots,m-1\}\}$ contains the configuration $\begin{Bmatrix} -1 & 1 \\ -1 & 1 \end{Bmatrix}$. We call this a *balancing configuration*. Now divide the square lattice into boxes such that (x,y) is in the box indexed by $(a,b), a,b \in \mathbb{Z}$ if $x \in \{ma, ma+1, \ldots, m(a+1)-1\}$ and $y \in \{mb, mb+1, \ldots, m(b+1)-1\}$. Call the box indexed by (a,b) 'open' if it contains the above balancing configuration, and 'closed' otherwise. These events are independent and have identical probability for all boxes. The fact that we have site percolation on the oriented lattice when boxes are open with probability greater than 3/4 now implies that with positive probability we have an infinite sequence of adjacent (that is, sharing one boundary edge) open boxes (see [14]). We now show that this event implies the existence of a just-visiting path that has bounded partial sums returning to 0 infinitely often.

Our path takes the shortest route from balancing configuration in one open box to balancing configuration in the next, in the infinite sequence of adjacent open boxes. When at a balancing configuration the path moves about on the -1s or 1s until its partial sum becomes zero. When this happens it goes to the next open box in the path. The partial sums never leave the interval $[-3m, 3m]$, so we have our path.

For the second part of the theorem we use the *enhanced square lattice*, which is obtained from the ordinary square lattice by adding edges between all pairs of vertices at Euclidean distance $\sqrt{2}$. We divide the square lattice

into disjoint boxes, $\{B(N+1)+((N+1)a,(N+1)b) : a,b \in \mathbb{Z}\}$. We call a box 'open' if it contains a vertex with a -1 assigned to it, and 'closed' otherwise. Now if there exists a circuit of adjacent closed boxes around the origin then there can be no path from the origin with partial sums bounded in some interval of length N. For large enough p (depending on N) the existence of this circuit follows from standard counting arguments. □

2.5 Asymptotics

Before we start with the involved proof of Theorem 2.2.4, it is worth explaining why the result is reasonable. When p is close to 1, there will be mostly $+1$s around with isolated -1s every now and then. Therefore, partial sums will typically increase. We can essentially only *decrease* partial sums along a just-visiting path, when we see two neighbouring vertices both with label -1. Therefore, these double -1s play an important role, and the only way to keep the partial sums bounded is to jump around from one such double -1 to the other. For p close to 1, the spatial distribution of these double -1s will be close to a Poisson process. Depending on the density of these double -1s it will or will not be possible for a sequence of partial sums to stay within a given bound. This is where continuum percolation comes in. If the double -1s are too far apart from each other, that is, if a certain percolation process does not percolate, then the bound will be large. If the percolation process does percolate, then we can use an infinite path in the percolation process to find a path with partial sums that remain bounded.

A significant amount of work is necessary to turn this idea into a proof. The connection with continuum percolation is made in Proposition 2.5.1, Lemma 2.5.2 and Lemma 2.5.6. In Lemma 2.5.3 and Lemma 2.5.7 this is translated into statements concerning the existence of paths with bounded partial sums for particular values of the parameter p. In the final proof of the theorem we then show how we can obtain the general result.

Fix $\lambda > 0$. On the grid $(\mathbb{Z}/N)^2$ assign the value 1 to each vertex with probability p_N, and the value -1 otherwise, independently of all other vertices, and where p_N is defined as to satisfy

$$2N^2(1-p_N)^2 = \lambda. \tag{2.5}$$

Next we put a point at the centre of every edge of the grid which has -1s at both ends, and the resulting points form a point process which we call Π_N. We shall also need the point process Π_N^+ defined as follows. Consider an edge which contains a point of Π_N. This point is in Π_N^+ if and only if the

parallel edge immediately to the right or above it has +1s at its endpoints. We call such a configuration of two adjacent −1s with two adjacent +1s to the right or above it a *balancing configuration*. Π_N^+ contains no other points, hence Π_N^+ is a subset of Π_N. We write $\Pi_N(S)$ to denote the number of points of Π_N in S. The number of points of Π_N in S that are due to edges that are completely contained in S is denoted by $\Pi_N(S^-)$. Similar definitions apply to Π_N^+.

Proposition 2.5.1 *The point processes Π_N converge weakly to a homogeneous Poisson process with rate λ when $N \to \infty$.*

To prove this more or less obvious result, it suffices to prove convergence of the avoidance function, that is, the probability that a finite union of rectangles contains no points, see [7], Proposition 9.1.IX. This can be done directly from the definition of the p_Ns. However, in the proof of the forthcoming Lemma 2.5.6 we shall need a estimate about weak convergence that is uniform for λ in a certain interval. For that estimate we shall use an explicit Stein-Chen upper bound on the total variation distance between two distributions. To save space later, we also use this Stein-Chen approach in the following proof of Proposition 2.5.1.

Proof of Proposition 2.5.1 Consider a union of rectangles R which, without loss of generality, we will assume has total area 1. We say that en edge e satisfies $e \in R$ if the centre of e is contained in R. We write G_e for the event that e has −1s on both endpoints. Now $\Pi_N(R)$ can be written as

$$\Pi_N(R) = \sum_{e \in R} \mathbf{1}_{G_e}.$$

where $\mathbf{1}_{G_e}$ denotes the indicator function of G_e. This indicator function is increasing in the number of −1s, and has expectation $(1 - p_N)^2$. We can therefore apply Corollary 2.E.1 in [4] and conclude that the total variation distance between a Poisson distribution with parameter $\mathbb{E}_p(\Pi_N(R))$ and $\Pi_N(R)$ is bounded above by

$$\frac{1 - e^{-\mathbb{E}_p(\Pi_N(R))}}{\mathbb{E}_p(\Pi_N(R))} \left(\mathrm{Var}(\Pi_N(R)) - \mathbb{E}_p(\Pi_N(R)) + 2 \sum_{e \in R} (1 - p_N)^4 \right). \qquad (2.6)$$

The asymptotic variance of $\Pi_N(R)$ is easily computed: $\Pi_N(R)$ is a sum of $(2 + o(1))N^2$ indicator random variables associated with the edges of the grid. When these edges are not adjacent, the corresponding indicator random variables are independent. Each edge (apart from those at the boundary

which have fewer) has 6 adjacent edges, and the probability that two adjacent edges both have a point from Π_N is equal to $(1-p_N)^3$. These observations lead to

$$\begin{aligned}\mathbb{E}_{p_N}(\Pi_N(R)^2) &= (2+o(1))N^2(1-p_N)2 + 6(2+o(1))N^2(1-p_N)^3 + \\ &\quad +(2+o(1))^2N^4(1-p_N)^4 \\ &\to \lambda+\lambda^2,\end{aligned}$$

when $N\to\infty$, because $\lambda = 2N^2(1-p_N)^2$. Clearly, $\lim_{N\to\infty}\mathbb{E}_p(\Pi_N(R)) = \lambda$. Hence $\mathrm{Var}(\Pi_N(R))$ converges to λ and since

$$\sum_{e\in R}(1-p_N)^4 = (2+o(1))N^2(1-p_N)^4 \to 0,$$

for $N\to\infty$, the bound (2.6) above implies that the total variation distance between the distribution of $\Pi_N(R)$ and a Poisson distribution with parameter $\mathbb{E}_p(\Pi_N(R))$ converges to 0. A Poisson distribution with parameter $\mathbb{E}_p(\Pi_N(R))$ certainly converges in distribution to a Poisson distribution with parameter λ, and as convergence in total variation implies convergence in distribution, we have that $\Pi_N(R)$ converges in distribution to a Poisson distribution with parameter λ. □

Next we centre a diamond with radius 1/2 around each point of Π_N. The result we shall call the *discrete diamond process*, to distinguish it from the *continuous diamond process* which is similar, but with diamonds centred around points of a Poisson process. The process obtained by centring diamonds around points of Π_N^+ is called the *reduced discrete diamond process*. Recall the definition of λ_c as the critical density associated with the continuous diamond process. The reason for the slightly larger radius in the next lemma will become apparent soon.

2.5.1 Proof of Theorem 2.2.4(a)

Lemma 2.5.2 *Let $\lambda<\lambda_c$. There exists $\epsilon>0$ such that for all $\delta>0$ we can find L sufficiently large such that for all large N, the probability (under $\mathbb{P}_{p_N}$) that the discrete diamond process with radii $\frac{1}{2}(1+\epsilon)$ connects $[0,3L]\times 0$ to $[0,3L]\times L$ within $[0,3L]\times[0,L]$ is at most δ.*

Proof Let $\lambda<\lambda_c$. Then using scaling relation (2.1) we can find $\epsilon>0$ such that $\lambda<\lambda_c(\frac{1}{2}(1+\epsilon))$. In words, after increasing the radii of the diamonds by a factor $1+\epsilon$, the continuous diamond process remains subcritical. Now let $\delta>0$. Since the continuous diamond process (with increased radii) is

subcritical, we have from Theorem 3.5 in [16] (stated for balls but the proof also works for diamonds) that the probability that this continuous diamond process connects $[0, 3L] \times 0$ to $[0, 3L] \times L$ is at most $\delta/2$, for L sufficiently large. We fix such an L. Choose ϵ' so small that if we partition $[0, 3L] \times [0, L]$ into squares with side length ϵ', with probability at least $1 - \delta/2$ we have that (i) a Poisson process with rate λ has at most one point in each of the squares; (ii) the connectivity structure of the diamonds (of radius $\frac{1}{2}(1 + \epsilon)$) around these points does not change when we move points of the Poisson process around in the squares with side length ϵ' in which they are contained.

We denote this set of squares with side length ϵ' by $\mathcal{S}$. Since Π_N converges weakly to a Poisson process with rate λ (Proposition 2.5.1), the $\mathbb{P}_{p_N}$-probability that a particular subset of $\mathcal{S}$ contains a point of Π_N, converges to the corresponding probability in a Poisson process with rate λ. Property (ii) above then guarantees that if the continuous diamond process does not connect opposite sides, neither does the discrete one. □

We next turn our previous lemma into a statement about paths with bounded partial sums.

Lemma 2.5.3 *For $\lambda < \lambda_c$, there exists N'' such that for all $N > N''$ there are $\mathbb{P}_{p_N}$-almost surely no just-visiting paths with partial sums bounded in $[0, N]$.*

For the proof of Lemma 2.5.3, we need some more propositions. The ϵ in the statements that follow is the ϵ dictated by λ in the statement of Lemma 2.5.2. If the discrete diamond process with increased radii does not connect the top and bottom sides of $[0, 3L] \times [0, L]$, we say that there is a *gap* in the long direction, or from $0 \times [0, L]$ to $3L \times [0, L]$. Other gaps are defined analogously. Lemma 2.5.2 says that the probability of a gap can be made as high as desired by taking L large (and large N of course).

Proposition 2.5.4 *For $\delta > 0$ and $\lambda < \lambda_c$ there exists L, such that for large N, with $\mathbb{P}_{p_N}$-probability at least $1 - \delta$ we can construct a (random) pair of curves, C_1 and C_2 contained in $[0, 3L] \times [0, L]$ such that*

1. *the curves are self-avoiding and made up of finitely many straight line segments,*

2. *the curves start in $0 \times [0, L]$ and end in $3L \times [0, L]$,*

3. *C_2 is the locus of points at distance $(1 + \epsilon)$ from C_1, above C_1,*

4. *the conditional (that is, given C_1 and C_2) joint distribution of the configuration between C_1 and C_2 is i.i.d. with the original marginals, conditioned on the event that there are no adjacent -1s.*

Proof According to the proof of Lemma 2.5.2 there exists L such that the probability of having an overlapping series of diamonds connecting top and bottom of the box $[0, 3L] \times [0, L]$ is smaller than $\delta/6$, and we fix a large such L. According to Lemma 2.5.2, we can now choose N' such that if we put a diamond of radius $\frac{1}{2}(1+\epsilon)$ at every pair of adjacent -1s, the probability that these connect top and bottom of $[0, 3L] \times [0, L]$ is smaller than $\delta/2$. Thus C_1 and C_2 with the first three properties above must exist with probability at least $1 - \delta/3$.

We next show that we can construct such a gap, without finding any more information about those 1s and -1s in the gap other than that there are no adjacent pairs of -1s, and possibly that some of the vertices on the edge of the gap have value 1. We will need to work on the extended rectangle $[-(1+\epsilon), 3L+(1+\epsilon)] \times [0, L]$, to take into account the dependencies between different areas. The probability that there is a gap crossing this extended rectangle is greater than $1 - \delta$ for large L. This can be seen by combining horizontal gaps in $[-(1+\epsilon), 3L-(1+\epsilon)]\times[0, L]$ and $[+(1+\epsilon), 3L+(1+\epsilon)]\times[0, L]$ with a vertical gap crossing $[0, L]^2$. Each of these gaps exists with probability at least $1 - \delta/3$, which means that the probability that they all exist must be at least $1 - \delta$. Any gap crossing this rectangle contains a gap crossing $[0, 3L] \times [0, L]$.

Order the edges in the extended rectangle $[-(1 + \epsilon), 3L + (1 + \epsilon)] \times [0, L]$, in some deterministic way and call this set E_L. Until the end of this proof all edges mentioned are assumed to be in E_L. If we have an edge e_a with -1s at both ends, we say that another edge e_b is in the cluster of e_a if there is a sequence of adjacent edges $(e_a = e_0, e_1, \ldots, e_k = e_b)$ all having -1s at both ends.

Set E_B^0 to be those (random) edges on the horizontal line closest to $[-(1+\epsilon), 3L + (1 + \epsilon)] \times 0$ with -1s at both ends, along with all the edges in the clusters containing these edges. Let E_G^0 be the set of those edges in E_L that lie on the same horizontal line and do not have -1s at both ends, along with any edges adjacent to those in E_B^0 yet not in it.

In an inductive fashion, to find E_G^{i+1} and E_B^{i+1} from E_G^i and E_B^i do the following. Take the first edge in E_G^i in the ordering, and check all edges in $E_L \cap (E_G^i \cup E_B^i)^c$ completely within distance $(1 + \epsilon)$ of this edge to see if they have -1s at both ends. If such an edge exists, set E_B^{i+1} to be E_B^i along with all such edges and any edges in their clusters around them. In addition, set E_G^{i+1} to be E_G^i along with any edges sharing a vertex with an

edge in E_B^{i+1}, yet not in it. In the case that the first edge in E_G^i has no edges in $E_L \cap (E_G^i \cup E_B^i)^c$ within distance $(1+\epsilon)$ with a -1 at both ends, move through the ordering of edges until either an edge is found that does have this property or there are no more edges in E_G^i. In the first case proceed as we would have done for the first edge, except with this edge, and in the latter case set E_G^{i+1} and E_B^{i+1} to be E_G^i and E_B^i.

There must be a value of i for which E_G^i and E_G^{i+1} are the same, as are E_B^i and E_B^{i+1}, if only because there are only finitely many vertices in the box. At this point either there is an edge in E_G^i which is closer than $(1+\epsilon)$ to the top of the box, or not. If this is so there can be no gap of size $(1+\epsilon)$ in the box by the manner of our construction, and if this is not so we must have found a gap (between the locus of points at least $(1+\epsilon)$ away from all the edges E_G and the points $(1+\epsilon)$ from this locus back towards the E_Gs). What do we know about the points in our gap? The ends of the edges in E_G closest to the edges in E_B must have value 1, but apart from this, our construction method has told us exactly that there are no pairs of -1s in the gap. □

Proposition 2.5.5 *Given $\delta > 0$ and $\lambda < \lambda_c$ there exists N'' and L such that, for all $N > N''$, the probability (under $\mathbb{P}_{p_N}$) that there exists a path with sums bounded in $[0, N]$ from $[0, 3L] \times 0$ to $[0, 3L] \times L$ in $[0, 3L] \times [0, L]$ is less than δ.*

Proof Take N, L so that the probability that a gap exists is at least $1 - \delta/2$. Then if there is a gap from $0 \times [0, L]$ to $3L \times [0, L]$, then there exist curves C_1 and C_2 as in Proposition 4. In order to have a path with sums bounded in $[0, N]$ from $[0, 3L] \times 0$ to $[0, 3L] \times L$ in $[0, 3L] \times [0, L]$, there must exist at least one path from C_1 to C_2 with partial sum bounded above by N. Since there are no adjacent -1s between C_1 and C_2, if such a path exists, then also a self-avoiding path from C_1 to C_2 with this property exists: indeed, in the absence of adjacent -1s, the final partial sum cannot be made less by adding loops. Denote the number of self-avoiding paths between C_1 and C_2 with partial sums bounded above by N by U_N. The standard FKG inequality (see for instance [9]) implies that, conditioned on the increasing event of having no adjacent -1s between the curves, the (conditional) probability of the event $\{U_N > 0\}$ is at most the unconditional probability of the same event, that is, its probability under $\mathbb{P}_{p_N}$. Note that such a path, of length M say, must contain at least $\frac{M-N}{2}$ -1s, if its partial sums are to be bounded above by N. Hence the conditional probability that $U_N > 0$ is bounded above by

$$\sum_{M \geq (1+\epsilon)N} 3L^2(N+1)^2 4 \cdot 3^{M-1} \sum_{j=(M-N)/2}^{M} \binom{M}{j} p_N^{M-j}(1-p_N)^j$$

$$\begin{aligned} &\leq \sum_{M \geq (1+\epsilon)N} 3L^2(N+1)^2 4 \cdot 3^{M-1}(1-p_N)^{\frac{M-N}{2}} 2^M \\ &= \frac{4L^2(N+1)^2(1-p_N)^{\epsilon N/2} 6^{(1+\epsilon)N}}{1-6(1-p_N)^{1/2}}. \end{aligned}$$

This converges to zero as $N \to \infty$. Thus we can find N' such that for $N > N'$ this is less than $\delta/2$. □

Proof of Lemma 2.5.3 First choose δ and then define the event that $(0,0)$ is *closed* as the occurrence of a gap in the long direction which is not crossed by a path with partial sums bounded in $[0,N]$, in each of the boxes $[0,3L]\times[0,L]$, $[0,L]\times[0,3L]$, $[2L,3L]\times[0,3L]$ and $[0,3L]\times[2L,3L]$. Note that by the previous proposition there exists N', L, so that this occurs with probability at least $1-4\delta$, for $N > N'$. We now ask whether this event occurs at other points, in particular at $\{(2aL,2bL) : a,b \in \mathbb{Z}\}$, in which case we say (a,b) is closed. These events are not independent, but the events '(a,b) is closed' and '(c,d) is closed' are independent if $(2aL-2cL)^2+(2bL-2dL)^2 > 8L^2$. This means we can compare this process to a 1-dependent ordinary site percolation model on the enhanced square lattice. A standard counting argument (see for instance [11]) shows that for δ sufficiently small, there can be no infinite self-avoiding path of non-closed points.

Any path with partial sums that are bounded in $[0,N]$ cannot pass through gaps without using a section of path that is itself bounded in $[0,N]$. Thus it must be contained in the non-closed areas, which are, by the above argument, almost surely finite. Therefore it cannot be infinite itself, almost surely. □

Proof of Theorem 2.2.4(a) If $\limsup_{N\to\infty} 2N^2(1-q_N)^2 < \lambda_c$, then we can find N_0, and λ such that for all $N > N_0$,

$$2N^2(1-q_N)^2 < \lambda < \lambda_c.$$

If we use this λ in Lemma 2.5.3, we see that there can be no paths with partial sums bounded in $[0,N]$ for N large enough, almost surely, for the sequence of p_Ns, defined by this λ. However for $N > N_0$, $p_N < q_N$, which means that we can couple the two discrete processes with parameters p_N and q_N in the natural way, that is, the set of vertices with value -1 under q_N is a subset of the corresponding set under p_N. Now it is a matter of carefully inspecting the proof of Lemma 2.5.3. We first showed that the limit point process Π_N is a Poisson process. The only place where this fact was used later was to assert that appropriate gaps exist in the subcritical Boolean model. The coupling just mentioned implies that point processes Π'_N associated with the q_Ns will be stochastically smaller than Π_N. (Note that we do not have information

about a possible weak limit of the Π'_Ns.) This means that gaps have an even higher probability to occur with the q_Ns than with the p_Ns. The estimate of the probability to bridge such a gap with bounded partial sums is monotone in p_N for N sufficiently large. □

2.5.2 Proof of Theorem 2.2.4(b)

The proof of the supercritical part of the theorem must be different as we will need to consider balancing configurations (recall the definition above) the occurrence of which, as they contain both 1s and −1s, is neither increasing nor decreasing in p. We begin by considering percolation of diamonds of the reduced discrete diamond process. In the next lemma, the uniformity of N_1 in λ will be important later on.

Lemma 2.5.6 *Given $\lambda_c < \lambda_{down} < \lambda_{up} < \infty$. There exists N_1 such that for all $N > N_1$ and λ such that $\lambda_{down} < \lambda < \lambda_{up}$ the reduced discrete diamond process percolates under $\mathbb{P}_{p_N}$.*

Proof of Lemma 2.5.6 We can choose $\epsilon > 0$ such that $\lambda_{down}(1-\epsilon)^3 > \lambda_c$. Rearranging and using scaling relation (2.1) above we find that, $\lambda_{down}(1-\epsilon) > \lambda_c(\frac{1}{2}(1-\epsilon))$. In words, after decreasing *both* the rate of the Poisson process and the radii of the diamonds by a factor $1-\epsilon$, the continuous diamond process remains supercritical. It is clear that this is also true (with the same ϵ) for any $\lambda > \lambda_{down}$.

Now consider the grid $(\frac{1}{4}\epsilon\mathbb{Z})^2$. A square of this grid is called a *grid square*. Any diamond of radius $\frac{1}{2}(1-\epsilon)$ centred in a grid square is then contained in any diamond of radius $\frac{1}{2}$ centred in that same grid square. For a grid square S, we denote by F_S the event that $\Pi_N^+(S^-) \geq 1$. We first claim that for $N \to \infty$,

$$\mathbb{P}_{p_N}(F_S) \to 1 - e^{-\lambda(\epsilon/4)^2}, \tag{2.7}$$

uniformly in $\lambda_{down} < \lambda < \lambda_{up}$. For this we use the Stein-Chen Poisson approximation anticipated in the proof of Proposition 2.5.1 to estimate the total variation distance between the distributions of the number of points in S from Π_N and that from a Poisson distribution of rate corresponding to the expected number of such points, $\mathbb{E}_p(\Pi_N(S))$. As in (2.6), we have that this total variation distance is at most

$$\frac{1-e^{-\mathbb{E}_p(\Pi_N(S))}}{\mathbb{E}_p(\Pi_N(S))}\left(\mathrm{Var}(\Pi_N(S)) - \mathbb{E}_p(\Pi_N(S)) + 2\sum_{\text{edges in } s}(1-p_N)^4\right).$$

A similar albeit somewhat more complicated computation as in the proof of Proposition 2.5.1 now yields that this is bounded above by

$$\frac{1-e^{-2(N+1)^2\left(\frac{\epsilon}{4}\right)^2(1-p_N)^2}}{2N^2\left(\frac{\epsilon}{4}\right)^2(1-p_N)^2}\times \left(6\cdot 2\left(\frac{\epsilon}{4}\right)^2(N+1)^2(1-p_N)^3+2\cdot 2\left(\frac{\epsilon}{4}\right)^2(N+1)^2(1-p_N)^4\right).$$

Keeping in mind relation (2.5), this then gives a bound which tends to zero uniformly in $\lambda_{down}<\lambda<\lambda_{up}$ when $N\to\infty$. The fact that

$$\left(\frac{\epsilon}{4}\right)^2\lambda\le\mathbb{E}_p(\Pi_N(S))\le\left(\frac{N+1}{N}\right)^2\left(\frac{\epsilon}{4}\right)^2\lambda$$

implies that the null probability of Poisson distribution of rate $\mathbb{E}_p(\Pi_N(S))$ converges to the null probability of a Poisson distribution of rate $\left(\frac{\epsilon}{4}\right)^2\lambda$. Claim (2.7) follows immediately.

Next we want to show that

$$\mathbb{P}_{p_N}(\{\Pi_N^+(S^-)=0\}\cap\{\Pi_N(S^-)>0\})\to 0, \tag{2.8}$$

uniformly in $\lambda_{down}<\lambda<\lambda_{up}$, as $N\to\infty$. The event in (2.8) can only occur if we see one of the configurations

$$\left\{\begin{array}{cc} & -1\\ -1 & -1\end{array}\right\},\left\{\begin{array}{cc} -1 & \\ -1 & -1\end{array}\right\},\left\{\begin{array}{cc} -1 & -1\\ -1 & \end{array}\right\}$$

somewhere in the grid square. The probability of this is at most $3(\epsilon N/4+1)^2(1-p_N)^3=(\epsilon N/4+1)^2(\frac{\lambda}{2})^{3/2}\frac{1}{N^3}$. This goes to zero uniformly for all $\lambda_{down}<\lambda<\lambda_{up}$, when $N\to\infty$.

According to (2.7) and (2.8), we find that uniformly in $\lambda_{down}<\lambda<\lambda_{up}$, for all N large enough,

$$\mathbb{P}_{p_N}(F_S)>1-e^{-\lambda(1-\epsilon)(\epsilon/4)^2},$$

which is the probability that there is a point in the grid square due a Poisson point from a Poisson process of rate $\lambda(1-\epsilon)$. Hence we can couple the continuous diamond process based on this Poisson process and the reduced discrete diamond process in such a way that whenever a grid square contains at least one Poisson point, it also contains at least one point of Π_N^+. The grid size was chosen in such a way that the union of the diamonds (of radius $\frac{1}{2}(1-\epsilon)$) of the Poisson points, are contained in the union of the diamonds

(with radius $\frac{1}{2}$) corresponding to Π_N^+. The former process is supercritical by the choice of ϵ, and therefore also the latter process is supercritical, which is what we wanted to prove. □

We next turn this lemma into a statement of paths with bounded partial sums.

Lemma 2.5.7 *Given $\lambda_c < \lambda_{down} < \lambda_{up} < \infty$ there exists N_1 such that for all $N > N_1$ and λ such that $\lambda_{down} < \lambda < \lambda_{up}$ there is a just-visiting path with partial sums bounded in $[0, N+1]$ with positive $\mathbb{P}_{p_N}$ probability.*

Proof of Lemma 2.5.7 Fix λ such that $\lambda_{down} < \lambda < \lambda_{up}$. Recall that $2N^2(1-p_N)^2 = \lambda$. Then, By Lemma 2.5.6, the origin is contained in an unbounded connected components of diamonds from the reduced discrete diamond process with positive probability. If this is the case, we have a self-avoiding path $\pi = (\pi_0 = 0, \pi_1, \ldots)$ on the grid $(\mathbb{Z}/N)^2$ starting at the origin such that π visits a balancing configuration at least every N steps. We may assume that if π visits a balancing configuration for the first time at time n_1, say, the partial sums up to that moment are all in $[0, N]$. Call this first-visited balancing configuration W_1. (It is possible that π_n is contained in the intersection of two balancing configurations in which case we just make a choice and call one of these W_1.) Let m_1 be the first time m after n_1 for which $\pi_m \notin W_1$. Define n_2 as $\min\{n \geq m_1 : \pi_n \text{ visits a balancing configuration}\}$. Note that it is possible that $n_2 = m_1$. The balancing configuration visited at time n_2 is denoted W_2, where it is again possible that we have a choice. Define n_k, m_k and W_k for $k = 1, 2, \ldots$ inductively in this fashion. We will now construct a just-visiting path π'. π' will be constructed from π by adding loops of repeating vertices each time π visits a balancing configuration. More precisely, π' follows π up to time n_1. Consider the sum of the labels along π until π visits the next balancing configuration W_2, that is, $S_k^\pi - S_{n_1}^\pi$, for $k = n_1 + 1, \ldots, n_2$. These numbers are uniformly bounded by N in absolute value. We construct the next part of π' as follows. Pass around in the balancing configuration W_1 until you have reached a sum which guarantees that when we after that travel to W_2 along π the partial sums will never be smaller than 0 or larger than $N+1$. When we are in W_2 we repeat this process. □

Proof of Theorem 2.2.4(b) We consider two possibilities:
(i) $\liminf_{N\to\infty} 2N^2(1-q_N)^2 = \infty$,
(ii) $\limsup_{N\to\infty} 2N^2(1-q_N)^2 < \infty$.

For case (i) we take a grid square S, and we simply note that the probability that S contains a balancing configuration converges to 1 as N tends

to infinity. Thus, with positive probability, there will be an infinite path of grid squares containing balancing configurations for N sufficiently large. Hence the diamonds of the reduced diamond process must percolate, and in the same way as in the proof of Lemma 2.5.7 we can move around between balancing configurations to find our required path.

In case (ii) we can find λ_{up}, λ_{down} such that for all large N,

$$\lambda_c < \lambda_{down} < 2N^2(1-q_{N-1})^2 q_{N-1}^2 < \lambda_{up} < \infty.$$

We can then define λ_N as $2N^2(1-q_{N-1})^2 q_{N-1}^2$ and apply Lemma 2.5.7 to this series of λ_Ns, which immediately gives the result. □

Proof of Corollary 2.2.5 For this corollary we will show that each of the following three cases leads to a contradiction:

(i) $\liminf_{p\to 1} 2(1-p)^2 I(p)^2 < \lambda_c$,

(ii) $\lambda_c < \limsup_{p\to 1} 2(1-p)^2 I(p)^2 < \infty$,

(iii) $\limsup_{p\to 1} 2(1-p)^2 I(p)^2 = \infty$.

In case (i) we can choose a sequence $p_1^\circ, p_2^\circ, \ldots$, such that $\lim_{i\to\infty} p_i^\circ = 1$ and such that $\limsup_{i\to\infty} 2(1-p_i^\circ)^2 I(p_i^\circ)^2 < \lambda_c$. Now as $\lim_{i\to\infty} I(p_i^\circ) = \infty$ the set $\{I(p_i^\circ);\, i = 1, 2, \ldots\}$ must be infinite. This means that we can next define a new sequence, $p_1^\bullet, p_2^\bullet, \ldots$, which is a 'more regular' version of $\{p_1^\circ, p_2^\circ, \ldots\}$, in such a way that we can apply Theorem 2.2.4(a). So, if $\{p_i^\circ : I(p_i^\circ) = N\}$ is non-empty, then we choose $p_N^\bullet$ to be any element of this set. If it is empty, we set $p_N^\bullet$ equal to 1.

Then $\limsup_{N\to\infty} 2(1-p_{N+1}^\bullet)^2 N^2 < \lambda_c$, so by Theorem 4(a) we can find N' such that for all $N > N'$, $\mathbb{P}p_N^\bullet$-almost surely, there are no paths with partial sums bounded in any interval of length N. However, from the construction we have that $I(p_N^\bullet) = N$ for infinitely many N, which contradicts the definition of I.

In case (ii) we proceed similarly but this time we use Theorem 2.2.4(b). Now we can choose a sequence $p_1^\dagger, p_2^\dagger, \ldots$, such that $\lim_{i\to\infty} p_i^\dagger = 1$, $\limsup_{i\to\infty} 2(1-p_i^\dagger)^2 I(p_i^\dagger)^2 < \infty$ and $\liminf_{i\to\infty} 2(1-p_i^\dagger)^2 I(p_i^\dagger)^2 > \lambda_c$. Again we define a new version of this sequence. If the set $\{p_i^\dagger : I(p_i^\dagger) = N\}$ is non-empty, then we define $p_N^\ddagger$ to be any element of this set. If the set is empty, we set $p_N^\ddagger$ equal to $1 - \sqrt{\frac{\lambda_c+1}{2N^2}}$. Here the term $1 - \sqrt{\frac{\lambda_c+1}{2N^2}}$ is chosen to make $2(1-p_N^\ddagger)^2 N^2 = \lambda_c + 1 > \lambda_c$.

Next note that

$$\limsup_{N\to\infty} 2(1-p_{N+1}^\ddagger)^2 (N-1)^2 < \infty$$

and that

$$\liminf_{N\to\infty} 2(1-p^{\ddagger}_{N+1})^2(N-1)^2 > \lambda_c.$$

Thus by Theorem 2.2.4(b) there exists N'' such that for all $N > N''$ there are, with positive probability, paths with sums bounded in an interval of length $N-1$. Thus, by its minimality, $I(p^{\ddagger}_N)$ must be less than or equal to $N-1$, for $N > N''$. However we know that $I(p^{\ddagger}_N) = N$, which gives the required contradiction.

Case (iii) is proved in the same way as case (ii), by considering in this case a sequence of ps for which $2(1-p)^2 I(p)^2$ tends to infinity in the limit. The only necessary modifications are to define $p^{\ddagger}_N$ as an arbitrary constant between 1/2 and 1 for cases where $\{p^{\dagger}_i : I(p^{\dagger}_i) = N\}$ is empty, and to use the other condition in Theorem 2.2.4(b). □

2.6 Recurrent sums

Proof of Theorem 2.2.6 We give a proof by contradiction, so we first assume that there is a value of $p > 1/2$ such that both,

$$\mathbb{P}_p(\exists \text{ a path } \pi \text{ for which } S^{\pi}_n \text{ does not converge to } +\infty) > 0, \tag{2.9}$$

and,

$$\mathbb{P}_p(\exists \text{ a path } \pi \text{ with } S^{\pi}_n = 0 \text{ i.o.}) = 0. \tag{2.10}$$

In the proof of Lemma 2.4.1 we showed that (2.10) implies that for all $z \in \mathbb{Z}$,

$$\mathbb{P}_p(\exists \text{ a path } \pi \text{ with } S^{\pi}_n = z \text{ i.o.}) = 0. \tag{2.11}$$

This and (2.9) imply that there exists with positive probability a path π^d with

$$\lim_{n\to\infty} S^{\pi^d}_n = -\infty.$$

For the construction we use the spiral π^u, shown in Figure 2.4, which covers all points in the square lattice, and is self-avoiding. We define a spiral circuit from a point, α, to be that sub-path of π^u, starting at α, which passes around the origin until at Euclidean distance $\sqrt{2}$ from α, see Figure 2.5. Since $p > 1/2$, we have that, with probability 1,

$$\lim_{n\to\infty} S^{\pi^u}_n = +\infty.$$

We next construct from π^d and π^u a new self-avoiding path with partial sums that return to zero infinitely often, by moving between the two paths in a

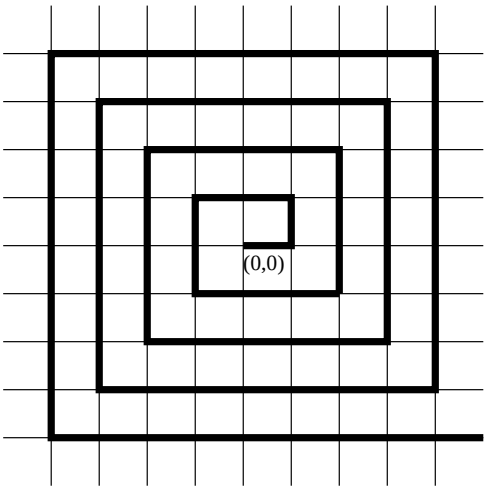

Figure 2.4: The spiral π^u

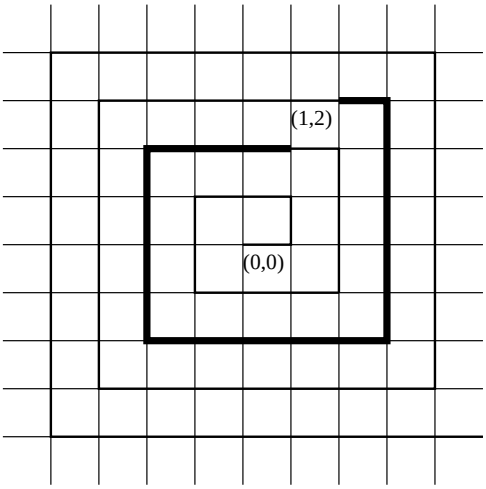

Figure 2.5: The spiral circuit from the point $(1, 2)$

way that ensures that the resulting path does not cross itself. This will then contradict (2.10).

We will speak of a point on a self-avoiding path as being *before* another (or as an *earlier* point) if it is nearer to the starting point, measuring along the path. Points that are *after* others (or *later* points) are defined analogously. We note two things:
(a) There are an infinite number of points on any self-avoiding path from the origin which have the property that no earlier points on the path intersect later points of the spiral.
(b) There exists a last intersection of a self-avoiding path and any spiral circuit, in the ordering given by the path.

We construct our new path π^n as follows, after starting by following (i.e. copying the steps of) π^d for a little while:

1. Follow π^d until a point when both the partial sum of π^n is below zero and none of the points on π^u after this point are already in π^n. This is possible by (a), and the fact that the partial sums of π^d are only above or at any value for a finite number of steps.

2. Follow π^u until a point when both the partial sum of π^n is above zero and none of the points on π^u after this point are already in π^n. Note that π^u will never return to earlier points in π^n, by the previous step.

3. Find the last intersection between π^d and the spiral circuit from the point where we currently are, and move along π^u to this point.

4. Repeat.

Notice that it is always possible to keep passing from path to path in this manner, and thus the partial sums of π^n are zero infinitely often. We have a contradiction, and so our result. □

Proof of Corollary 2.2.7 The existence of paths with partial sums which return to zero infinitely often is clearly a tail event and so has probability 1 or 0, by Kolmogorov's 0-1 law.

Assign to each vertex $v \in \mathbb{Z}^2$ a uniform $[0,1]$ random variable, U_v. We can use these variables to couple realisations at all parameter values together such that if $\begin{cases} U_v \geq 1-p & \text{then} \quad X_v = 1, \\ U_v < 1-p & \text{then} \quad X_v = -1. \end{cases}$

Theorem 2.2.6 states that if with positive probability there is a path with partial sums not converging to $+\infty$, then there exists with positive probability, and hence with probability one, a path with partial sums that return to zero infinitely often. On the other hand, if all paths have partial

sums converging to $+\infty$, then clearly there does not exist a path with partial sums that return to zero infinitely often. Hence we have a path with partial sums that return to zero infinitely often, with positive probability, if and only if there is a path, with positive probability, whose partial sums do not converge to $+\infty$. From the coupling described above we see immediately that the probability of the event that all paths have partial sums converging to $+\infty$ is monotone in p. It follows that the probability of having a path with partial sums that return to zero infinitely often is also monotone in p. □

Acknowledgements We thank Deepak Dhar and Jeffrey Steif for various suggestions and fruitful discussions.

Bibliography

[1] S. E. Alm. (1993). Upper bounds for the connective constant of self-avoiding walks. *Combinatorics, Probability and Computing*, 2(2), 115–136.

[2] M. J. Appel and J. C. Wierman (1987). Infinite AB clusters exist on the triangular lattice. *Journal of Physics A: Mathematical and General*, 20, 2533–2537.

[3] M. J. Appel and J. C. Wierman (1987). On the absence of infinite AB clusters in bipartite graphs. *Journal of Physics A: Mathematical and General*, 20, 2527–2531.

[4] A. Barbour, L. Holst, and S. Janson (1992). *Poisson Approximation*. Oxford Science Publications.

[5] I. Benjamini and H. Kesten (1995). Percolation of arbitrary words in $\{0,1\}^{\mathbb{N}}$. *Annals of Probability*, 23(3), 1024–1060.

[6] I. Benjamini and Y. Peres (1994). Markov chains indexed by trees. *Annals of Probability*, 22(1), 219–243.

[7] D. Daley and D. Vere-Jones (1980). *An Introduction to the Theory of Point Processes*. Springer.

[8] K. J. Falconer (1990). *Fractal Geometry: Mathematical Foundations and Applications*. Wiley.

[9] G. R. Grimmett (1999). *Percolation*. Springer, second edition.

[10] G. R. Grimmett and D. Stirzaker (1992). *Probability and Random Processes*. Oxford University Press, second edition.

[11] H. Kesten (1980). *Percolation for Mathematicians*. Birkhauser.

[12] H. Kesten and V. Sidoravicius and Y. Zhang (1998) Almost all words are seen in critical site percolation on the triangular lattice. *Electronic Journal of Probability*, 3(10), 1–75.

[13] H. Kesten and V. Sidoravicius and Y. Zhang (2001) Percolation of arbitrary words on the close-packed graph of $\mathbb{Z}^2$. *Electronic Journal of Probability*, 6(4), 1–27.

[14] T. M. Liggett (1995). Survival of discrete time growth models, with applications to oriented percolation. *Annals of Applied Probability*, 5(3), 613–636.

[15] T. Łuczak and J. C. Wierman (1989). Counterexamples in AB percolation. *Journal of Physics A: Mathematical and General*, 22, 185–191.

[16] R. Meester and R. Roy (1996). *Continuum Percolation*. Cambridge University Press.

[17] M. V. Menshikov and S. A. Zuev (1993). Models of ρ-percolation. Probabilistic Methods in Discrete Mathematics (Petrozavodsk, 1992). *Progress in Pure and Applied Discrete Mathematics*, 1, 337–347. VSP, Utrecht.

Chapter 3

Covering algorithms, continuum percolation and the geometry of wireless networks

Lorna Booth, Jehoshua Bruck, Massimo Franceschetti and Ronald Meester

Continuum percolation models in which each point of a two-dimensional Poisson point process is the centre of a disc of given (or random) radius r, have been extensively studied. In this paper, we consider the generalisation in which a deterministic algorithm (given the points of the point process) places the discs on the plane, in such a way that each disc covers at least one point of the point process and that each point is covered by at least one disc. This gives a model for wireless communication networks, which was the original motivation to study this class of problems.

We look at the percolation properties of this generalised model, showing that an unbounded connected component of discs does not exist, almost surely, for small values of of the density λ of the Poisson point process, for any covering algorithm. In general, it turns out not to be true that unbounded connected components arise when λ is taken sufficiently high. However, we identify some large families of covering algorithms, for which such an unbounded component does arise for large values of λ.

We show how a simple scaling operation can change the percolation properties of the model, leading to the almost sure exis-

tence of an unbounded connected component for large values of λ, for any covering algorithm.

Finally, we show that a large class of covering algorithms, that arise in many practical applications, can get arbitrarily close to achieving a minimal density of covering discs. We also show (constructively) the existence of algorithms that achieve this minimal density.

3.1 Introduction and motivation

Geometric covering algorithms have been extensively studied in the last 20 years, in the context of computational geometry and combinatorial optimisation (see the survey by Agarwal and Sharir (1998), section 7.1). More recently, distributed versions of these algorithms have been proposed in the context of wireless network architectures (see Gerla and Tsai 1995).

Continuum percolation models (also referred to in the literature as Poisson Boolean models) were introduced by Gilbert (1961) to model wireless networks of radio transmitting stations, and they have been extensively investigated by mathematicians since then. In these models discs of a given (or random) radius r are centred at each point of a two-dimensional Poisson process X. The a.s. (almost sure) existence of unbounded connected components of discs, for a given density λ of the point process is often considered.

We consider the generalisation in which a deterministic algorithm (given the points of the point process) places the discs on the plane, in such a way that each disc covers at least one point of X, and each point is covered by at least one disc (see Figure 3.1 for a visual example).

Our aim is twofold: on one side we explore the mathematics of the new model and we answer some very natural questions that arise from a pure mathematical point of view. On the other side, we note that most of our results can be applied to rigorously model geometric properties of wireless communication networks.

Random graphs are the natural tool that is often used to model communication networks. In such graphs vertices represent communication endpoints and edges represent two-way channels. In the standard model of Erdős and Rényi (1959, 1960, 1961a,b), each pair of vertices has some probability (the same for all pairs of vertices, regardless of their separation) of being joined by an edge. Therefore, for each natural number n, there is a a probability space consisting of all graphs with exactly n vertices. Erdős and Rényi proved that many interesting properties of random graphs occur a.s. as $n \to \infty$. Their model of random graphs, however, is not suited to accurately repre-

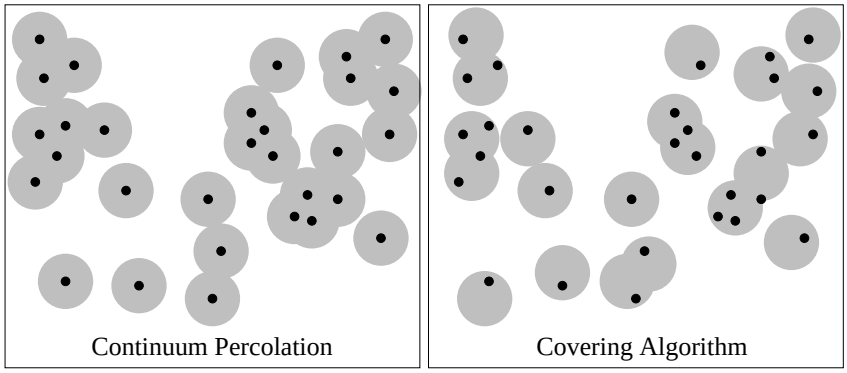

Figure 3.1: **Two different disc coverings of a random point process.** *A continuum percolation model places a disc at each random point; in our generalised model a covering algorithm places (possibly fewer) discs to cover all the points, according to a deterministic rule. We are interested in the almost sure existence of an unbounded connected component of discs, for a given density of points* λ.

sent networks of short-range radio transmitting stations. This motivated Gilbert (1961) to propose an alternative model in which the range of the transmitters is a parameter. In his paper, he constructed a random network by considering a two-dimensional Poisson point process and joining each pair of points by an edge if discs of radius r centred at those two points intersect. He was the first to introduce the concept of continuum percolation, identifying a phase transition behaviour, i.e., the existence of a critical value λ_c for the density of the Poisson point process, at which an unbounded connected graph a.s. forms and the network can provide some long distance communication. His results were later extended, from a purely mathematical standpoint, by Hall (1985), Menshikov (1986), Roy (1990), Meester and Roy (1994), and others, leading to a theory of random coverage processes (see the books by Hall (1988) and Meester and Roy (1996)). In a more applied framework, Gupta and Kumar (1998, 2000) and Penrose (1997) recently used a similar model to determine the throughput capacity of a wireless network.

Our model of communication refines the one introduced by Gilbert, considering a wireless backbone that routes data packets through the network. In our model, we differentiate between base stations and clients: clients communicate between each other by connecting to base stations that forward their messages to their destinations (see Figure 3.2). If a client is within a given distance of a base station, it can connect to it and we say that the client is covered by the base station. Hence, the network appears as a set of circular cells (base stations broadcast domains) that cover a set of points (clients). A covering algorithm decides where to place the cells, according to the distri-

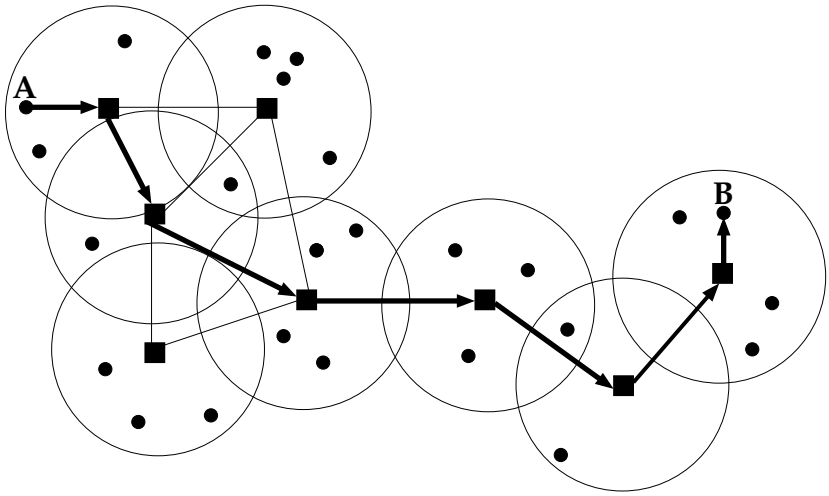

Figure 3.2: **The Wireless Backbone.** *A connected component of discs forms the wireless backbone of the network. If a client A wants to communicate to a client B, it connects to the closest base station and its message is routed through the backbone in a multi-hop fashion, until it reaches client B.*

bution of the clients. The algorithm can be a distributed, self-organising one, in a model where the entire population of clients elects 'cluster heads' and divides itself into subsets that are covered by the cluster heads (see Gerla and Tsai, 1995); or a more centralised one, in a model where the clients are mobile and the base stations are static. In the latter case, the base stations could be laid on a fixed grid and the covering algorithm could determine the subset of them that need to be turned on, at any given time, to provide coverage (see Franceschetti, Cook, and Bruck 2001). The algorithm would typically try to minimise the number of base stations that need to be turned on to cover all the clients, or, in the dynamic case, would try to minimise the base stations movement needed to cover all the clients.

As in Gilbert's model, we assume a completely wireless network, hence, base stations can connect to other base stations only up to a limited distance. We assume two base stations to be connected only if the corresponding discs overlap (although we will modify this requirement later in the paper). Therefore, if two clients are in the same connected component of overlapping discs, they can communicate, because they are reached by a connected path of base stations of that component.

In percolation theory one is interested in unbounded connected components. In our setting, unbounded connected components are of interest because they represent long-range communication. The almost fully connected

state of the wireless network is also interesting, and any of the proof techniques we use to prove percolation will also show that most of the clients fall inside the unbounded connected component and are therefore able to communicate, under similar conditions. In this case, a few additional 'bridge' stations can be added to connect isolated components to the unbounded one and achieve the full connectivity of the network.

It may be argued that not allowing transmitters to be further than a certain distance from a client is an artificial constraint if we are interested in having long-range communication. However we wish to see when long-range communication occurs spontaneously, given that the clients require communication with the base stations and the natural restriction of not placing a base station where it will not be used.

We informally summarise our main results in the next section; in Section 3 we introduce some notation and definitions; Sections 4 is devoted to existence and non-existence results for unbounded components, for different covering algorithms; Section 5 considers the effect of varying the base stations communication radius; Section 6 discusses existence of optimal covering algorithms; Section 7 contains some open problems.

3.2 Summary of results

Our results can be grouped into four categories: non-existence results, existence results for different classes of covering algorithms, scaling results, and results concerning the optimality of certain algorithms. In the following, we let X be a two-dimensional Poisson point process of density λ. The points of X represent the clients that are covered by base stations.

Non-existence results. Our first results regard the non-existence of an unbounded connected component of covering discs. We show (Theorem 3.4.1) that for any algorithm covering all the points of X by discs of radius r, there exists a $\lambda_0 > 0$ such that for all $0 < \lambda \leq \lambda_0$, P_λ(there is an infinite component) $= 0$. Then we show that the symmetric result, i.e., the a.s. existence of an unbounded connected component for large values of λ, is not generally true, but depends on the type of covering algorithm. It is known that a covering that places a disc centred at each point of X forms a.s. an unbounded connected component for large values of λ (Gilbert 1961). In order to show that this result does not generalise to all coverings, we specify a covering algorithm that does not form an unbounded connected component for any value of λ.

Existence results. We proceed by identifying different families of covering

algorithms that form an unbounded connected component a.s. for large values of λ. One of such coverings, that is practical for our applications, is a *grid covering*. We show (Theorem 3.4.3) that for any algorithm covering all the points of X by discs of radius r centred at the vertices of a grid, there exists a $\lambda_1 < \infty$, such that for all $\lambda > \lambda_1$, P_λ(there is an infinite connected component) $= 1$.

Another family of coverings that we consider are the *flat coverings*. A flat covering has the property that the restriction to any box of size $n \times n$ contains at most $k(n)$ discs, for any value of λ. We show in Theorem 3.4.4 that there exists a $\lambda_1 < \infty$, such that for all $\lambda > \lambda_1$, P_λ(there is an infinite connected component) $= 1$.

A third class of algorithms that we examine are the *shift invariant coverings*. These algorithms are defined by the requirement that the covering commutes with shifts of the points. We ask whether a shift invariant algorithm necessarily forms a.s. an unbounded connected component for large values of λ, and we answer this question negatively, by constructing a shift invariant algorithm that does not exhibit this property.

Scaling results. We then introduce a further extension of our model. We note that when we consider overlapping discs as connected components, then we implicitly assume, in our model of a wireless network, that the maximum radius of communication between two base stations is twice as large as the maximum radius of communication between clients and base stations. This observation leads to the natural question of what would happen if the ratio between the two radii is different from two.

In the standard Poisson Boolean model, that places a disc centred at each point of a Poisson point process X, considering a different radius for connectivity corresponds to a simple scaling operation, hence it does not change the basic properties of the model. In our extended model, however, this leads to more interesting results. Call r the radius of the discs used to cover the points of X and R the maximum distance sufficient to connect disc centres. We show (Theorem 3.5.1) that:

- If $R/r \leq 1$, then, for any grid G, there is a covering algorithm that places discs only at the vertices of G, and a.s. does not form an unbounded connected component, for any value of λ.

- If $1 < R/r < 2$, then, for some given dense grid G, there is a covering algorithm that places discs only at the vertices of G, and a.s. does not form an unbounded connected component, for any value of λ.

- If $R/r = 2$, then, for any grid G, any covering algorithm that places

discs only at the vertices of G forms a.s. an unbounded connected component for large values of λ.

- If $R/r > 2$, then any algorithm forms a.s. an unbounded connected component for large values of λ, even if it is not grid-based.

Note that the latter case is useful in practice, because it states that if base stations can communicate at a distance larger than twice the maximum communication distance to the clients, an unbounded connected component forms a.s. for large values of the density of the clients, regardless of the covering algorithm used to build the cellular network.

Optimality Results. Finally, we show (constructively, in Theorem 3.6.3) the existence of algorithms that are optimal in achieving a minimal density of covering discs. We also show that a certain class of practical algorithms can achieve densities arbitrarily close to the optimal.

3.3 Notation and definitions

Let $\mathbb{R}^2$ be the Euclidean plane, let $\mathcal{B}^2$ be the σ-algebra of Borel sets in $\mathbb{R}^2$ and let $\ell(\cdot)$ be Lebesgue measure in $\mathbb{R}^2$. Let N be the collection of all counting measures on $(\mathbb{R}^2, \mathcal{B}^2)$, which assign finite measure to bounded Borel sets and for which the measure of a point is at most 1. In this way, N can be identified with the set of all configurations of points in $\mathbb{R}^2$, without limit points. Let $\mathcal{N}$ be the σ-algebra of N generated by sets of the form $\{\nu \in N : \nu(A) = k\}$, for all integers k and bounded Borel sets A. A (planar) point process X is defined as a measurable mapping from a probability space $(\Omega, \mathcal{F}, P)$ into $(N, \mathcal{N})$. For $A \in \mathcal{B}^2$, we denote by $X(A)$ the random number of points inside A. In this paper, X will always be a Poisson process with density $\lambda > 0$. We sometimes abuse notation and write $x \in \nu$, for $x \in \mathbb{R}^2$ and $\nu \in N$, to express that x is one of the points of ν.

We define a shift operation $T_t : \mathbb{R}^2 \to \mathbb{R}^2$ as a translation in $\mathbb{R}^2$ over the vector $t \in \mathbb{R}^2$, such that $T_t(x) = t + x$ for all $x \in \mathbb{R}^2$. The shift T_t induces in a natural way a shift transformation on N, which we also denote by T_t. Let, for all $x \in \mathbb{R}^2$ and $r \geq 0$, $D(x, r)$ be the disc of radius r centred at x: $D(x, r) = \{y \in \mathbb{R}^2 : |y - x| \leq r\}$. A circle of radius r centred at x is the set $\{y \in \mathbb{R}^2 : |y - x| = r\}$. The boundary of a set A will be denoted by ∂A.

We call two discs D_i, D_j *adjacent* if $D_i \bigcap D_j \neq \emptyset$. We write $D_i \leftrightarrow D_j$ if there exists a sequence $D_{i_1}, D_{i_2}, \ldots, D_{i_k}$ of discs such that $D_{i_1} = D_i$, $D_{i_k} = D_j$, and D_{i_l} is adjacent to $D_{i_{l+1}}$ for $1 \leq l < k$. A *(connected) component* or *cluster* is a set $\{D_i : i \in J\}$ of discs which is maximal with the property that

$D_i \leftrightarrow D_j$ for all $i, j \in J$. We identify a component with the set of centres of the discs in it.

We next formally define a covering algorithm: A *covering algorithm* $\mathcal{A}$ with discs of radius r, is a measurable mapping $\mathcal{A} : N \to N$ with the following properties:

1. for all $x \in \mathcal{A}(\nu)$ there exists $y \in \nu$ such that $y \in D(x, r)$,
2. for all $y \in \nu$ there exists $x \in \mathcal{A}(\nu)$ such that $y \in D(x, r)$.

We define the *occupied region* C of $\mathcal{A}(\nu)$ as the union $\bigcup_{x \in \mathcal{A}(\nu)} D(x, r)$.

In this paper, we examine different classes of covering algorithms, which we define as follows:

1. *Grid Algorithms.* Let $G \subset \mathbb{R}^2$ be the set of all vertices of a two dimensional lattice. A grid algorithm $\mathcal{A}$ constrains the covering discs to be centred at the vertices of G. That is, $x \in \mathcal{A}(\nu)$ implies $x \in G$. Naturally we require G to be dense enough that every point can be covered by a disc centred at some vertex.
2. *Flat Algorithms.* A flat algorithm $\mathcal{A}$ has the property that its restriction to any box of size $n \times n$ contains at most $k(n)$ discs, for some $k(n) < \infty$.
3. *Finite Horizon Algorithms.* Let $B_n(x)$ be the box of size $n \times n$ centred at x, and let, for all $\nu \in N$, $\nu|_{B_n(x)}$ denote the restriction of ν to $B_n(x)$. In other words, $\nu|_{B_n(x)}$ can be identified with the set of points $\{\nu \cap B_n(x)\}$. We say that a covering algorithm $\mathcal{A}$ has finite horizon if there exists a constant $h \geq 0$ (the *horizon*), so that whenever $\nu|_{B_{n+2h}(x)} = \nu'|_{B_{n+2h}(x)}$, we have $\mathcal{A}(\nu)|_{B_n(x)} = \mathcal{A}(\nu')|_{B_n(x)}$, for all n and x. In words, this means that changing ν outside $B_{n+2h}(x)$ does not change the covering inside $B_n(x)$.
4. *Shift Invariant Algorithms.* A shift invariant algorithm $\mathcal{A}$ is defined by the property that $T_t(\mathcal{A}(\nu)) = \mathcal{A}(T_t(\nu))$, for all t. In words, this means that the covering algorithm commutes with shifts of the points.
5. *n-Square Algorithms.* An n-square algorithm is obtained as follows. Partition the plane into boxes of size $n \times n$. For each such box B_n, the covering of the points inside B_n should use the minimal number of discs possible.

Suppose now that we want to cover the points of X by the covering algorithm $\mathcal{A}$, that is, we consider the measurable map $\mathcal{A} \circ X : \Omega \to N$. This Boolean model is denoted by $(X, \mathcal{A}) = (X, \lambda, r, \mathcal{A})$, where λ is the density of

X, and r the radius of the covering discs. The law of this process is denoted by $P_{\lambda,r}$. The standard Poisson Boolean model that places a disc of radius r, centred at each point of X is obtained when we take $\mathcal{A}$ to be the identity, and is denoted by (X, λ, r). In this model there exists $\lambda_c(r)$ such that for $\lambda \leq \lambda_c(r)$ we have no infinite cluster a.s., while for $\lambda > \lambda_c(r)$ there is an infinite cluster with probability 1. We often denote $\lambda_c(1)$ by λ_c and scaling implies that $\lambda_c(r) = \lambda_c(1)/r^2$ (see Meester and Roy (1996) for more details).

Next, we define the *density* of $(X, \mathcal{A})$. Let $N_{(X,\mathcal{A})}(n)$ be the (random) number of discs centred inside the box $B_n(0)$. The density of $(X, \mathcal{A})$ is given by

$$\lim_{n\to\infty} N_{(X,\mathcal{A})}(n)/n^2,$$

whenever this limit (i) exists a.s. and (ii) is an a.s. constant.

Finally, we introduce one more piece of terminology. If $(X, \mathcal{A})$ contains an unbounded component of discs with positive probability, we say that $(X, \mathcal{A})$ *percolates*.

3.4 Percolation

In this long section we think of r as being fixed, while λ varies. Accordingly, we sometimes write $P_\lambda = P_{\lambda,r}$. We also use P to mean $P_{1,1}$. The expectation under P we denote by E.

Our first result deals with the lack of percolation for small values of λ.

Theorem 3.4.1 *For any covering algorithm $\mathcal{A}$, there exists a $\lambda_0(r) > 0$ such that for all $0 < \lambda \leq \lambda_0$, $(X, \lambda, r, \mathcal{A})$ does not percolate.*

Proof of Theorem 3.4.1. Assume that, with positive probability, there is an unbounded connected component of covering discs for $(X, \lambda, r, \mathcal{A})$. Then with positive probability, there is an unbounded connected component in the Poisson Boolean model $(X, \lambda, 2r)$. That is because two intersecting covering discs in $(X, \lambda, r, \mathcal{A})$ cover points that are at a distance of at most $4r$ to each other; and the Poisson Boolean model $(X, \lambda, 2r)$ places discs of radius $2r$ at each of the covered points. We then choose $\lambda_0 = \lambda_c/(2r)^2$, so that $(X, \lambda, 2r)$ does not form an unbounded connected component a.s. for $\lambda \leq \lambda_0$. □

A symmetric result to Theorem 3.4.1, i.e., percolation for large values of λ, depends on the type of covering algorithm used:

Proposition 3.4.2 *There exists a covering algorithm $\mathcal{A}$, such that for all λ, $(X, \lambda, r, \mathcal{A})$ does not percolate.*

Proof of Proposition 3.4.2. The proof is constructive. Draw circles of radii $\{3kr, k \in \mathbb{N}\}$ around the origin, and notice that a.s. no Poisson point falls on any of these circles. Then cover the Poisson points, with discs of radius r, without intersecting these circles. Notice that the circles divide the plane into finite annuli and, since each cluster of discs resides in at most one of these finite annuli, each cluster must be bounded, whatever the value of λ. □

We next look at families of algorithms that do percolate for large values of λ, beginning by considering grid algorithms.

Theorem 3.4.3 *For any grid covering algorithm $\mathcal{A}$, there exists a $\lambda_1 < \infty$, such that $(X, \lambda, r, \mathcal{A})$ percolates for all $\lambda > \lambda_1$.*

This theorem can be proved by application of Theorem 3.4.4, as all grid algorithms are flat. The proof below however is more elegant and offers more insight into the structure of grid algorithms.

Proof of Theorem 3.4.3. The proof relies on a construction that maps the covering discs to a discrete site-percolation model. We illustrate the idea by considering a square lattice and a distance between two neighbouring lattice vertices of one. Call a covering disc centred at a lattice vertex a grid disc. Clearly, the radius of a disc must be $r \geq \frac{\sqrt{2}}{2}$, in order to be able to cover all possible points on the plane by using only grid discs. For any $r \geq \frac{\sqrt{2}}{2}$, the number of grid discs that intersect a lattice square $ABCD$ is finite and they partition the square into some number k_r of small regions A_i (see Figure 3.3). If at least one point of the Poisson point process falls into each region A_i, then the entire square $ABCD$ must be covered by grid discs. Now view each lattice square as a site of a site percolation model. Call the site occupied if there is at least one point of the Poisson process situated inside each region A_i, for $i = 1 \dots k_r$. Note that the occupancy of a site is independent of the occupancy of other sites and the probability of a site being occupied is given by: $p = \prod_{i=1}^{k_r} (1 - e^{-\lambda \ell(A_i)})$. Moreover, if two adjacent sites are both occupied, then the corresponding covering discs form a connected component. Thus, if there is an unbounded component of occupied adjacent sites, then there is an unbounded connected component of covering discs. Next, we choose λ large enough so that $p > p_c$, where p_c is the critical probability for site percolation on a square lattice. The a.s. existence of an unbounded connected component of covering discs immediately follows. □

Next, we consider flat algorithms. Recall that each such algorithm, $\mathcal{A}$, has the property that the restriction of $\mathcal{A}$ to any box of size $n \times n$ contains at

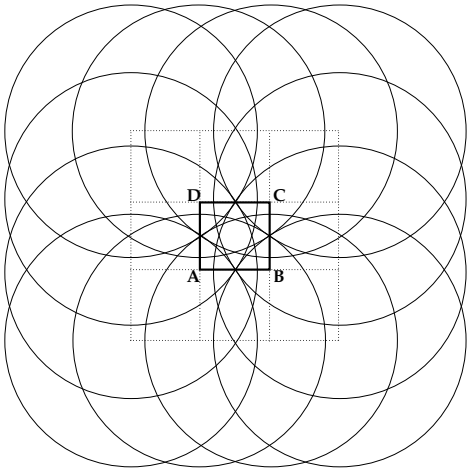

Figure 3.3: **Mapping to the site percolation model.** *The grid discs partition the square ABCD into a finite number of small areas.*

most $k = k(n)$ discs, for any value of λ. Note that this really is a weak requirement, since we can *completely* cover the box using at most $\alpha\lceil(n/r)\rceil^2$ discs, for some $\alpha \leq 1$. Any 'sensible' algorithm should therefore satisfy the requirement. Note that $k(n) < \infty$ for some n, immediately implies that $k(m) < \infty$, for any m, as we can cover an $m \times m$ square by a finite number of $n \times n$ squares.

Theorem 3.4.4 *Let $\mathcal{A}$ be a flat covering algorithm. Then there exists $\lambda_1 < \infty$, so that $(X, \lambda, r, \mathcal{A})$ percolates for all $\lambda > \lambda_1$.*

At first sight, the statement of the theorem is counterintuitive, since we claim that we force percolation by *restricting* the number of discs. The point is that by restricting the number of discs (independently of λ), the requirement of covering all points with this restricted number of discs makes percolation unavoidable.

Another version of this theorem requires an upper bound on the density of discs, the shift invariance of the algorithm under a pair of linearly independent shifts and the algorithm to be finite horizon.

Theorem 3.4.5 *Let $\mathcal{A}$ be a finite horizon covering algorithm such that a constant $\delta < \infty$ exists such that*

$$\limsup_{n\to\infty} \frac{\textit{number of discs in } B_n}{n^2} < \delta, \ \textit{a.s.} \ ,$$

and which is stationary under a pair of linearly independent shifts. Then there exists $\lambda_1 < \infty$, so that $(X, \lambda, r, \mathcal{A})$ percolates for all $\lambda > \lambda_1$.

Johan Segers has pointed out that although finite horizon algorithms of bounded density which are stationary under any *pair* of linearly independent vectors must percolate, for a sufficiently high density of points, there exist finite horizon covering algorithms of bounded density which are stationary under shifts of *one* vector, and do not percolate for any λ.

Before we prove Theorems 3.4.4 and 3.4.5, we first state and prove a preliminary geometric lemma.

Lemma 3.4.6 *Consider a collection of discs of radius r, with the property that at most $k(n) < \infty$ discs intersect any box of size $n \times n$. Then there exists an $\epsilon = \epsilon(n, r) > 0$ with the following property: if there are, either, at least two clusters that intersect the boundaries of both $B_{n+r}(x)$ and $B_{n+2r}(x)$, or a cluster wholly contained in $B_{n+2r}(x)$, then there is a disc of radius ϵ, contained in $B_{n+3r}(x)$, which is not intersected by any disc.*

Proof of Lemma 3.4.6. We write $B_n = B_n(x)$. All discs that intersect B_{n+3r} must be centred inside B_{n+5r}. Therefore, at most $k = k(n + 5r)$ discs intersect B_{n+3r}. Let C be a component that intersects the boundaries of both B_{n+r} and B_{n+2r}. The number of discs in C that intersect B_{n+2r} is denoted by l. Note that $l \leq k$.

Consider a section AB of the perimeter of C, from the boundary of B_{n+r} to the boundary of B_{n+2r}, which does not intersect either of these boundaries except at its ends (see Figure 3.5). This section has length at least $r/2$, and consists of parts of the boundaries of at most l discs, each of which appears only once. This latter fact follows from the observation that, since the distance between the boundaries of the two boxes is only $r/2$, any disc that contributes to an arc in AB, must overlap the boundary of at least one box. Moreover, note that were AB to contain two disjoint arcs from the same disc, then any disc overlapping that disc in order to make these arcs disjoint, must overlap the boundary of the box not overlapped by the first disc. Therefore these arcs would be in disjoint sections of $\partial C \cap B_{n+2r}/(B_{n+r} \cup \partial B_{n+2r})$. See Figure 3.4. It follows that at least one arc in AB is of length at least $r/2l$. Call this arc a.

Note that there are at most $k - 1$ discs intersecting B_{n+3r}, if we do not count the one that has a as a part of its boundary, and none of them intersect a, except at its end points. If we divide a into k arcs of equal size, then each of these discs will be nearest to one of these smaller arcs - assign this arc to this disc. One of the smaller arcs (of size at least $r/2kl$) will, however, have no disc is assigned ot it. This means that the space left by discs tangent to the ends of this smaller arc cannot be covered, and we can choose ϵ so small that a disc of radius ϵ fits into this space (see Figure 3.5). The value of ϵ that we have to choose only depends upon r and n.

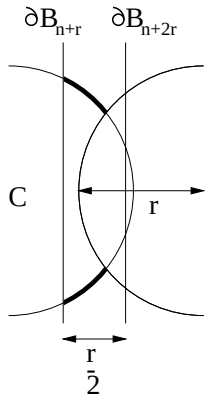

Figure 3.4: **No discs can have two disjoint components in AB.** Two disjoint arcs from the perimeter of one disc, in a component C connecting the two boundaries of B_{n+r} and B_{n+2r}, must be in disjoint parts of $\partial C \cap B_{n+2r}/(B_{n+r} \cap \partial B_{n+2r})$.

The same argument applies to a component wholly contained in B_{n+2r}, by considering its perimeter rather than its boundary between B_{n+r} and B_{n+2r}. □

Proof of Theorem 3.4.4. Let $t, u \in \mathbb{Z}$, and denote the box of size $n \times n$ centred at (tn, un) by $B_n(tn, un)$, as before. Let ϵ be chosen as in Lemma 3.4.6. We say that the vertex (t, u) is a *neighbour* of (t', u') if the boxes $B_n(tn, un)$ and $B_n(t'n, u'n)$ share an edge or corner. We call a vertex (t, u) *good* if all discs of radius ϵ contained in $B_{n+3r}(tn, un)$ contain at least one point of the Poisson process. Denote the event that (t, u) is good by $G(t, u)$. It is clear that when $\lambda \to \infty$, the probability of $G(t, u)$ converges to 1. It is also clear that $G(t, u)$ and $G(t', u')$ are independent whenever $\max\{|t - t'|, |u - u'|\} \geq \lceil \frac{(3r)}{n} \rceil$, for n larger than $2r$. Hence, the configuration of good sites is formed through a discrete, finite-range dependent percolation process, and it follows then from Durrett and Griffeath (1983) that for λ high enough, the good vertices percolate, i.e., contain an infinite component of good squares with probability one. What does this mean for our covering? Consider a good square B_n. By Lemma 3.4.6 any component cannot be wholly contained in B_{n+2r}, therefore, a component that covers points inside B_{n+2r}, must also intersect the boundary of B_{n+2r}. Also by Lemma 3.4.6, there can be only one component that intersects the boundaries of both B_{n+2r} and B_{n+r}. For n larger than 2ϵ such component exists and must reach to within 2ϵ of *all* edges of B_{n+r}, as, by the definition of a good square, there is no disc of radius ϵ inside B_{n+r} without any points of the Poisson process inside it (see Figure 3.6). Thus, the components associated with adjacent

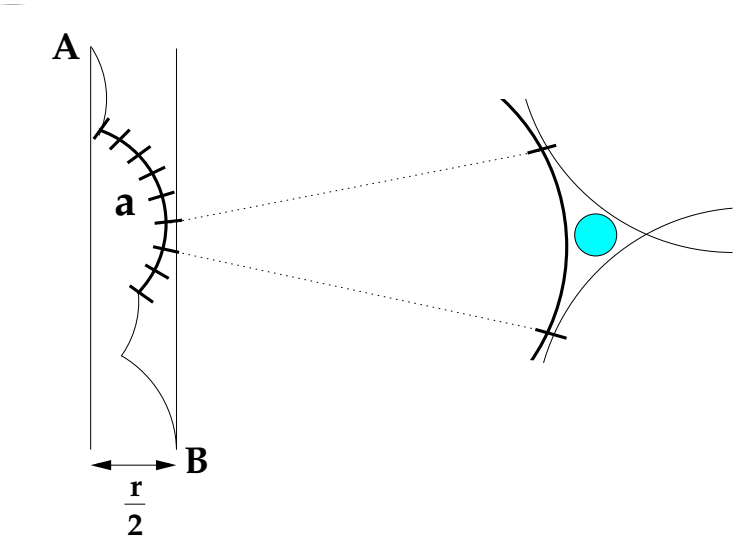

Figure 3.5: **Subdividing the arc.** An edge AB of a component C connecting the two boundaries of B_{n+r} and B_{n+2r}, has length at least $r/2$. This edge contains an arc a of length at least $r/2l$. Arc a is divided up into k sections, and by one of these we can place a small disc of radius ϵ that is not contained in any cluster.

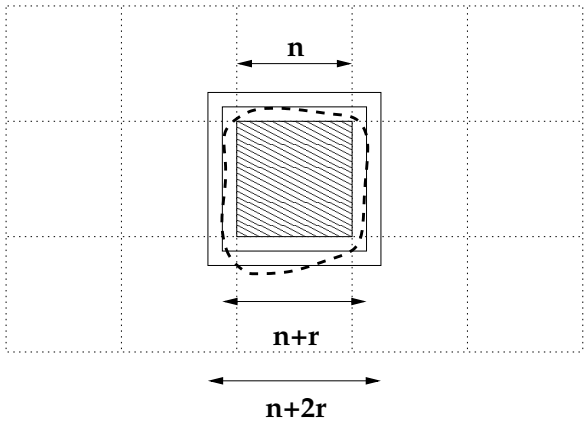

Figure 3.6: **A good square.** There is only one component of discs (represented by the dashed line) that intersects both B_{n+2r} and B_{n+r}. This component must reach to within 2ϵ of all edges of B_{n+r}, and will therefore intersect a component of an adjacent good square.

good squares must overlap, and we must have an infinite component of discs with probability one. □

We thank Johan Segers for the remark that this proof does not depend upon the algorithm being finite horizon.

Proof of Theorem 3.4.5.

The proof in this case is a little more complicated, but uses the same idea as that of Theorem 3.4.4. Without loss of generality assume that the algorithm be stationary under the shifts $T_{(m,0)}$ and $T_{(0,m)}$, for some m. Let h be the horizon of the algorithm, and assume $m > \max(1, h + 5r)$, again without loss of generality. Choose $\gamma > 0$ sufficiently small that $1-\gamma$ is strictly above the critical point for site percolation on the lattice $\{(tm, tu) : t, u \in \mathbb{Z}\}$ with edges between neighbouring sites. Take n, a multiple of m, so large that the probability that B_{n+5r} is intersected by more than $(\delta+1)(n+5r)^2$ discs is less than γ, uniformly in λ. We then use $(\delta+1)(n+5r)^2$ as our k in the lemma, and find an ϵ such that, if we have at most $(\delta+1)(n+5r)^2$ discs intersecting B_{n+5r} in the way described in the lemma, then we must have a disc of radius ϵ empty of Poisson points.

In the proof of the theorem we then call B_n good if B_{n+3r} has both no disc of radius ϵ empty of Poisson points in it and B_{n+5r} contains at most $(\delta+1)(n+5r)^2$ points. Other boxes are called good analogously. If λ is high enough, these boxes percolate, and we again have an infinite component of discs. □

We have seen that both finite-horizon, shift-invariant algorithms under a bounded density condition and flat algorithms necessarily percolate for high enough λ. It is natural to ask whether this always holds for shift invariant algorithms. It turns out that for these algorithms, large values of the density λ of the points do not guarantee the a.s. existence of an unbounded connected component. This is shown by describing a shift invariant covering algorithm that does not form an unbounded connected component for all λ, thus proving the following theorem:

Theorem 3.4.7 *There exists a shift invariant covering algorithm $\mathcal{A}$ of all the points of X by discs of radius r, such that for all λ, $(X, \lambda, r, \mathcal{A})$ does not percolate.*

The proof of Theorem 3.4.7 is constructive and rather technical. The covering we describe will have density λ.

Without loss of generality, consider covering discs of radius $r = 1$. The main idea is the following: given a realization of the Poisson point process, we

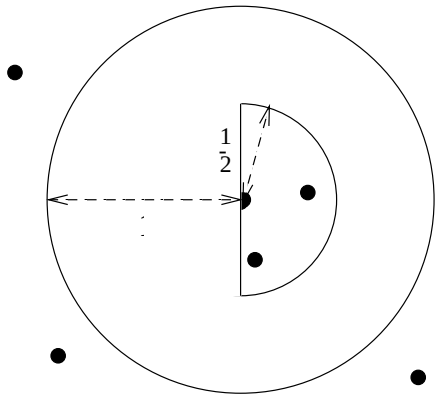

Figure 3.7: **A potential point** A potential point is a Poisson point with at least one other point in the half-disc of radius $\frac{1}{2}$ to the right of it, and no points in the disc of radius 1 centred at it, except in that half-disc.

first build a large structure of circles similar to those obtained in continuum fractal percolation models. We do this by placing circles of radii 18^m, with $m \in \mathbb{N}$, where we see certain configurations of points in the plane. The resulting structure is composed of clusters that are finite, but contain every bounded region of the plane. We then derive a shift invariant covering of all the points of X by discs of radius $r = 1$, leaving an empty space near to the boundaries of these clusters.

We illustrate the proof taking the density λ to be 1. The proof for the covering of a Poisson process of another density follows in the same way (in the proofs to follow, only the values of the ϵ's change).

We define a *potential-point* to be a Poisson point with at least one other point in the half-disc of radius $\frac{1}{2}$ to the right of it, and no points in the disc of radius 1 centred at it, except in the aforementioned half-disc (see Figure 3.7). Notice that the potential points cannot come within distance 1 of each other.

Given a decreasing sequence of positive numbers, $b_1 = \frac{1}{2}, b_2, b_3 \ldots$, an *m-point*, for $m \in \mathbb{N}$, is a potential-point which has its nearest neighbouring point between b_m and b_{m+1} away.

We start by proving a few lemmas.

Lemma 3.4.8 *For $\epsilon > 0$ sufficiently small, there exists a sequence $b_1 = \frac{1}{2}, b_2, b_3, \ldots$, such that the density of m-points is exactly $\epsilon 18^{-2m}$, for each $m \in \mathbb{N}$.*

Proof of Lemma 3.4.8. The density of potential points, λ_p, is calculable.

Choose $\epsilon > 0$ so that

$$\sum_{m=1}^{\infty} \epsilon 18^{-2m} = \epsilon \frac{1}{18^2 - 1} \leq \lambda_p.$$

We can now define b_i inductively, i.e. given b_m we choose b_{m+1} so that the density of m-points is exactly $\epsilon 18^{-2m}$. □

We consider a circle of radius 18^m around every m-point. Call such a circle an *m-circle*.

Lemma 3.4.9 *Every bounded region of the plane is a.s. wholly contained in some m-circle, for some m.*

We will use Proposition 7.3 of Meester and Roy (1994). This states:

Proposition 3.4.10 (Meester and Roy, 1994) *Let S be a stationary point process in $\mathbb{R}^d$, and let ρ be a non-negative random variable. If $\mathbb{E}(\rho^d) = \infty$, then in the Boolean model (S, ρ) the occupied component is a.s. $\mathbb{R}^d$.*

Furthermore the proof of the theorem can be adapted to show that, under the conditions of the theorem, every bounded region of the plane is a.s. wholly contained in some circle.

Proof of Lemma 3.4.9. Let S be the random collection of m-points, for all m. S is stationary. Note that the radii associated with points in S are independent, as the points of S do not come within distance 1 of each other. Let ρ be the radius distribution of the circles. Then (S, ρ) is a Boolean model, and the occupied component corresponds to all areas contained in some circle. In addition

$$\mathbb{E}(\rho^2) = \sum_{m=1}^{\infty} (18^m)^2 \epsilon 18^{-2m} = \infty,$$

so we can apply the proof of Proposition 3.4.10 and conclude that every bounded region is contained in some circle.

□

Now that we have shown that our circles a.s. contain any bounded region, we want to show that any cluster of intersecting circles is a.s. finite. To do this, we give a slight variation of a proof for fractal percolation that is in the book by Meester and Roy (1996) (Theorem 8.1). Again, we proceed by proving a series of lemmas but first we need a couple of definitions.

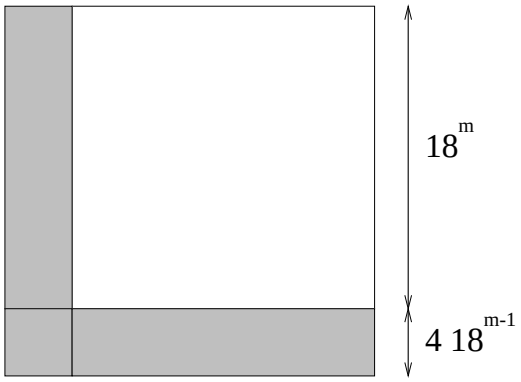

Figure 3.8: **Sets of possible dependence.**

We define the *sets of possible dependence* to the square $[0, 18^m]^2$ as the rectangles $(I_1 \times I_2 : I_i \in \{[-4 \times 18^{m-1}, 0], [0, 18^m]\}) \backslash [0, 18^m]^2$ (see Figure 3.8). Sets of possible dependence to other squares of the same size are the natural translations of this. Call a union of sets of possible dependence to a certain square, a *known region*. Define A_m to be the number of m-points in the square $[0, 18^m]^2$, and $A_{>m}$ to be the number of k-points, with $m < k \in \mathbb{N}$, within distance 4 of the square. Let C_m^J be the number of m-points in the known region J to the square.

Lemma 3.4.11 *For any $\delta > 0$, we can find m' and, uniformly in $m > m'$, $\epsilon > 0$, sufficiently small that:*

$$P(A_m > 0 | A_{>m} = 0 \cap C_m^J = 0) \leq \delta,$$

for any known region, J, to the square $[0, 18^m]^2$.

Proof of Lemma 3.4.11. We are interested in

$$\begin{aligned} P(A_m > 0 | A_{>m} = 0 \cap C_m^J = 0) &= \frac{P(A_m > 0 \cap A_{>m} = 0 \cap C_m^J = 0)}{P(A_{>m} = 0 \cap C_m^J = 0)} \\ &\leq \frac{P(A_m > 0)}{P(A_{>m} = 0 \cap C_m^J = 0)} \\ &\leq \frac{P(A_m > 0)}{1 - P(A_{>m} > 0) - P(C_m^J > 0)}, \end{aligned}$$

if $P(A_{>m} > 0) + P(C_m^J > 0) < 1$. We recall that, for a non-negative integer-

valued random variable N, $P(N > 0) \leq E(N)$, obtaining:

$$\begin{aligned} P(A_m > 0 | A_{>m} = 0 \cap C_m^J = 0) &\leq \frac{P(A_m > 0)}{1 - E(A_{>m}) - E(C_m^J)} \\ &\leq \frac{E(A_m)}{1 - E(A_{>m}) - E(C_m^J)}. \end{aligned}$$

if $E(A_{>m}) + E(C_m^J) < 1$.

Now, $E(A_{>m}) = (18^{2m} + 4(18^m + 1))18^{-2m}\frac{\epsilon}{18^2-1}$, which is the area of the region within distance 4 of the square multiplied by the total density of k-points, for all $k > m, k \in \mathbb{N}$. Noting that the maximum area of the known region is $(18^m(1 + 4/18))^2 - (18^m)^2$, we see that

$$E(C_m^J) \leq ((18^m(1 + 4/18))^2 - (18^m)^2)\epsilon 18^{-2m}.$$

The bound becomes:
$P(A_m > 0 | A_{>m} = 0 \cap C_m^J = 0) \leq$

$$\frac{\epsilon}{1 - (18^{2m} + 4(18^m + 1))18^{-2m}\frac{\epsilon}{18^2-1} - ((18^m(1 + 4/18))^2 - (18^m)^2)\epsilon 18^{-2m}},$$

which we can make uniformly less than δ by choosing ϵ small enough. □

Let $H(18^n) = [-\frac{3}{2} \times 18^n, \frac{3}{2} \times 18^n]^2 \backslash (-\frac{1}{2}18^n, \frac{1}{2}18^n)^2$. Define $I_{H(18^n)}$ to be the maximal connected cluster of (possibly partial) circles and the boundary of the box $[-\frac{1}{2}18^n, \frac{1}{2}18^n]^2$, fully contained in $H(18^n)$, and define $O_{H(18^n)}$ to be the maximal connected cluster of circles and the boundary of $[-\frac{3}{2} \times 18^n, \frac{3}{2} \times 18^n]^2$, fully contained in $H(18^n)$. See Figure 3.9.

Let $G(18^n)$ be the event that there is a *gap* in $H(18^n)$, i.e. the minimal distance between $I_{H(18^n)}$ and $O_{H(18^n)}$ is at least 18.

Lemma 3.4.12 *For $\epsilon > 0$ sufficiently small,*

$$\lim_{n \to \infty} P(G(18^n)) \geq \frac{1}{2}.$$

We prove Lemma 3.4.12 in two parts. Let $G_1(18^n)$ be the event that there is a gap between $I_{H(18^n)}$ and $O_{H(18^n)}$ in $H(18^n)$, when we consider circles only of radius 18^{n-1} or less. The size 18^{n-1} is chosen because it is a convenient size comparable to the size of $H(18^n)$. Let $G_2(18^n)$ be the event that no circles of radius 18^n or greater intersects $H(18^n)$ at all. Clearly, if $G_1(18^n)$ and $G_2(18^n)$ both occur, then $G(18^n)$ does also. Thus, Lemma 3.4.12 follows from the following two lemmas.

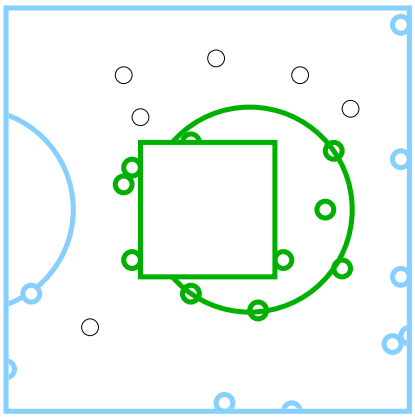

Figure 3.9: **The sets $I_{H(18^n)}$ and $O_{H(18^n)}$.** $I_{H(18^n)}$ is shown in dark grey and $O_{H(18^n)}$ in light grey.

Lemma 3.4.13 *For $\epsilon > 0$ sufficiently small,*

$$\lim_{n\to\infty} P(G_1(18^n)) \geq \frac{3}{4}.$$

Lemma 3.4.14 *For $\epsilon > 0$ sufficiently small,*

$$\lim_{n\to\infty} P(G_2(18^n)) \geq \frac{3}{4}.$$

Proof of Lemma 3.4.13. This proof closely follows that of Theorem 8.1 of Meester and Roy (1996) for fractal percolation, except for a number of extra technicalities. We are going to show that we dominate a version of the process which has more independence.

We first divide $H(18^n)$ into 8×18^2 sub-squares of size 18^{n-1} in the obvious way. We call two squares of the same size neighbours if they share an edge or corner.

Suppose for a moment that the probability that a square of size 18^m contains an m-point, is uniformly δ, for all m, independently of the occurrence of k-points anywhere, with $k \neq m, k \in \mathbb{N}$, and independently of the occurrence of m-points outside the square. We give our proof initially under this assumption, and then compare our original process with this.

We consider the 8×18^2sub-squares of $H(18^n)$, and in an order such that for any two squares, B and $B^{'}$ say, B is considered before $B^{'}$ if B is neither

to the right of nor above B'. We examine, in this order, these sub-squares of size 18^{n-1}, looking for $(n-1)$-points, in the following inductive fashion:

- to begin, all squares are declared to be neither corrupt nor bad.
- if a sub-square is not corrupt, then we examine the whole of it, looking for $(n-1)$-points. If it contains any, then we call it *bad*, and its neighbouring squares of the same level *corrupt*.
- if a sub-square is corrupt then we do not examine it.

Squares are bad if they contain centres of circles of comparable size to themselves, which then may extend into the corrupt squares. Corrupt squares may or may not contain $(n-1)$-points. We are careful not to find out this information, as it might tell us something about the distribution of points in the squares we have not yet considered.

We can then divide up each of the *good* squares (those that are neither bad nor corrupt) into 18^2 pieces, obtaining at most $8 \times 18^{2\times 2}$ squares of size 18^{n-2}, and we examine those, in an order such those squares nearer the bottom-left hand corner are considered first, looking for $(n-2)$-points, in the same inductive fashion as above. We end up declaring each of the squares of size 18^{n-2} that are sub-squares of good squares in $H(18^n)$ to be good, bad or corrupt.

We divide up each of the good squares of size 18^{n-2} into 18^2 squares of size 18^{n-3}, and use the same procedure to declare each good, bad or corrupt. We can then divide up each of the good squares, and repeat this procedure, while we still have good squares, and to a minimum square size of 18.

We then work backwards through the squares, starting with the smallest, to declare each either *dreadful* or not. A square of size 18 is dreadful if it is bad. In an inductive fashion, a square of size 18^m is dreadful if it is either a) bad or b) good but contains 2 or more dreadful squares of size 18^{m-1}. We call $H(18^n)$ dreadful if it contains any dreadful squares of size 18^{n-1}.

Under our temporary independence assumption, a square of size 18 is dreadful with probability δ, conditioned on the fact that it has not been declared corrupt before being checked. Thus, the probability that it is dreadful is at most δ. Then the probability p_m that a square of size 18^m is dreadful, is the probability it is a) bad or b) good but contains 2 or more dreadful squares of size 18^{m-1}. The probability that it is bad, is, as for a square of size 18, at most δ. The probability that it is good is at most $1-\delta$. It may be that some of the sub-squares of size 18^{m-1} of this square are corrupt, due to being neighbours of bad squares of the same size outside this square. Let N

be the number of such squares. As the probability of a square being dreadful is maximal when $N = 0$, it follows that:

$$p_m \leq \delta + (1-\delta)((1-p_{m-1})^{18^2} - 18^2 p_{m-1}(1-p_{m-1})^{18^2-1}) .$$

Letting

$$f(p,\delta) = \delta + (1-\delta)(1-(1-p)^{18^2} - 18^2 p(1-p)^{18^2-1}),$$

this becomes $p_m \leq f(p_{m-1}, \delta)$. Note that $p_1 = \delta$. If we can now show, for all δ, that

$$0 \leq p \leq b(\delta) \text{ implies } 0 \leq f(p,\delta) \leq b(\delta),$$

for some $b(\delta) \geq \delta$, it will follow that $p_m \leq b(\delta)$ for all m. We need $b(\delta)$ to be a bound that tends to 0 with δ.

Note that $f(p,\delta)$ is continuous in p and δ, that $f(0,\delta) = \delta$ and that $\frac{\partial f}{\partial p}(0,\delta) = 0$. It follows that $f(p,\delta) = p$ has a solution in $[0,1]$ for δ sufficiently small. The smallest such solution we call $b(\delta)$. Notice that $b(\delta) \geq \delta$, as $f(0,\delta) = \delta$ and $\frac{\partial f}{\partial p}(p,\delta) \geq 0$ on $[0,1]$. $\frac{\partial f}{\partial p}(p,\delta) \geq 0$ on $[0,1]$ also tells us that $0 \leq p \leq b(\delta)$ implies $0 \leq f(p,\delta) \leq b(\delta)$. Since $f(0,0) = 0$, $\frac{\partial f}{\partial p}(0,\delta) = 0$ and f is continuous, $\lim_{\delta \downarrow 0} b(\delta) = 0$.

We have shown that the probability that a square of size 18^m is dreadful is bounded from above, uniformly in m, by a function of δ that we can make arbitrarily small by choosing δ small enough. It then follows that the probability that $H(18^n)$ is dreadful (i.e. contains any dreadful squares of size 18^{n-1}) can be made as small as we like by choosing δ sufficiently small. We choose δ so that this probability is less than $\frac{1}{4}$.

We can now come back to our original process, and give up the independence assumption. We make the following comments:

- The probability that a square of size 18^m contains an m-point, when we come to check it, given any of the information we already have found, is at most δ, by Lemma 3.4.11, and because we never consider a corrupt square.

- By choosing ϵ sufficiently small we can make the bound δ as small as we need.

Let us consider what it means for $H(18^n)$ not to be dreadful. We argue that, in this case, we cannot have a connection by circles, of the appropriate sizes, from the inside to the outside of the box.

We first note that any m-point in $H(18^n)$ is either in a bad or a corrupt box of size 18^m. An m-circle may thus only intersect a box of size 18^m, if

that box is either bad, corrupt, or neighbours a corrupt square of the same size, unless it neighbours the boundary of $H(18^n)$.

We now give a series of definitions. Call a box of size 18^m *dodgy* if it is either bad, corrupt, neighbours a corrupt square of the same size, is dreadful, neighbours a dreadful square, or neighbours the border of $H(18^n)$. Call it *clean* if it is not dodgy. Call two boxes of size 18^m adjacent if they share an edge. Define an *m-circuit* to be a series of boxes of size 18^m, $(B_1, B_2, B_3, \ldots, B_k)$ where B_i is adjacent to B_{i+1} for $i = 1, 2, \ldots, k-1$ and B_k is adjacent to B_1. We also require that the circuit cuts off the origin from infinity.

A sub-circuit of boxes of size 18^j,$(B_1^j, B_2^j, B_3^j, \ldots, B_{k_j}^j)$, of a circuit of boxes of size 18^m (with $j < m$),$(B_1^m, B_2^m, B_3^m, \ldots, B_{k_m}^m)$, is a circuit of boxes of size 18^j inside the circuit of boxes of size 18^m, such that there exists $0 = i_1, i_2, i_3, \ldots, i_{k_m-1}, i_{k_m} = k_j$, such that $B_{i_l+1}^j, B_{i_l+2}^j, B_{i_l+3}^j, \ldots, B_{i_{l+1}}^j$ are contained in B_l^m. This means that the first few boxes of size 18^j are contained in B_1^m, the next few are in B_2^m, and so on.

We say that the property E_m holds if any m-circuit consisting of clean boxes of size 18^m contains a sub-circuit of boxes of size 18 that are not intersected by k-circles, for all $k \leq m$. Our objective is to show that E_m holds for $m = 1, 2, \ldots, n-1$, by induction on m.

If $m = 1$, we have a 1-circuit of clean boxes. Therefore there can be no 1-circle intersecting any of these boxes, and E_1 holds.

The inductive hypothesis is that E_{m-1} holds. In order to perform the induction step and prove that E_m also holds, consider an m-circuit of clean boxes of size 18^m, as depicted by in Figure 3.10. In this figure, the smallest squares are of size 18^{m-1}. If we show the existence of a sub-circuit of clean boxes of size 18^{m-1} inside the m-circuit, then, by the inductive hypothesis, the occurrence of E_m will follow.

By construction, all the boxes of size 18^m in our circuit are clean. This tells us that none of them intersects an m-circle. It also means that there is at most one dreadful square inside each box of size 18^m in the circuit. These are depicted in Figure 3.10 as black squares. A dreadful square may cause a 5×5 block of squares of size 18^{m-1} to be dodgy (grey 5×5 blocks in Figure 3.10). This can happen because the dreadful square can neighbour a corrupt square, which also has neighbours. Finally, any of the squares of size 18^{m-1} neighbouring the edge of the circuit, or the neighbours of these squares, may be dodgy, due to the proximity of dreadful squares just outside the circuit. This latter case is depicted in Figure 3.10 by the two grey circuits of width $2 \times 18^{m-1}$, along the edges of the m-circuit. These are all the possibilities for dodgy squares of size 18^{m-1} in our circuit.

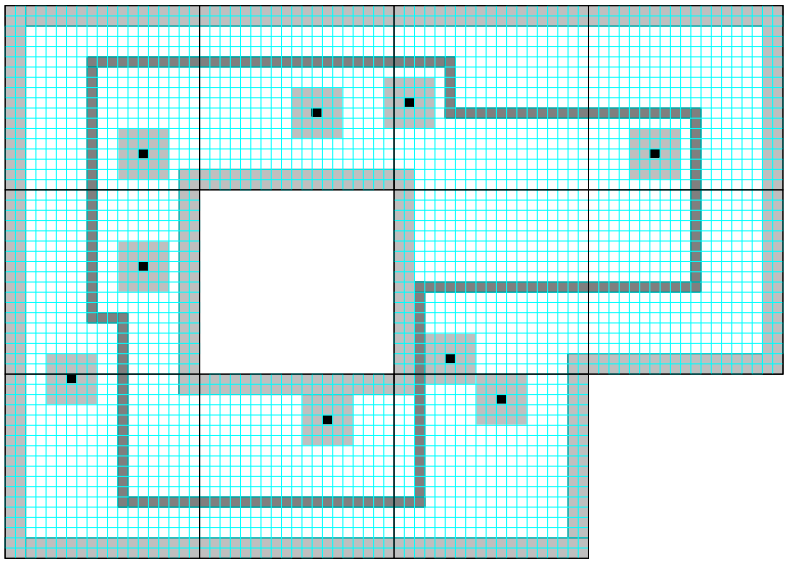

Figure 3.10: **Avoiding dodgy squares.** A circuit of squares of size 18^m contains a circuit of squares of size 18^{m-1}, avoiding dodgy squares.

By considering all possible arrangements of the dodgy squares we see that there must be a circuit of squares of size 18^{m-1} inside our circuit that avoids those dodgy squares (see Figure 3.10). Then, by the inductive hypothesis, there must also be a sub-circuit of squares of size 18, so E_m holds.

H_{18^n} is a circuit of boxes of size 18^n, and if all are clean, by the argument above, it follows that we must have a sub-circuit of boxes of size 18 that are not intersected by any m-circles, for $m = 1, 2, \ldots, n$, which is what we wanted to prove. □

Proof of Lemma 3.4.14. Let X be the number of circles of radius at least 18^n that intersect $[-\frac{3}{2} \times 18^n, \frac{3}{2} \times 18^n]^2$. A circle can only intersect the boundary of the box if the distance between the centre of the circle and the centre of the box is within $3 \times 18^n \times \frac{1}{\sqrt{2}}$ of the circle's radius (see Figure 3.11). Now write:

$$\begin{aligned} P(X > 0) &\leq E(X) \\ &\leq \sum_{b=n}^{\infty} \epsilon 18^{-2b} \pi((18^b + \frac{3 \times 18^n}{\sqrt{2}})^2 - (18^b - \frac{3 \times 18^n}{\sqrt{2}})^2). \end{aligned}$$

We can make this less than $1/4$ for all n by our choice of ϵ, implying that the probability of a large circle intersecting the box is less than $1/4$. □

We can now prove the finiteness of our clusters. We define a *thickened*

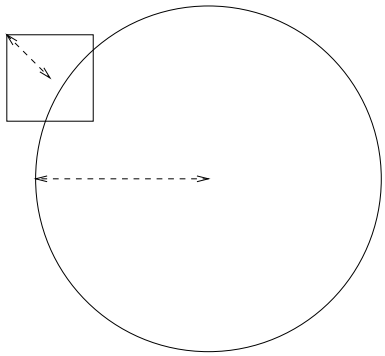

Figure 3.11: **Intersection between circle and box.** A circle of radius at least 18^n can touch the boundary of the box $[-\frac{3}{2} \times 18^n, \frac{3}{2} \times 18^n]^2$ only if the distance between the centre of the circle and the centre of the box is within $3 \times 18^n \times \frac{1}{\sqrt{2}}$ of the circle's radius.

cluster as a maximal connected component of points strictly within distance 9 of any circle.

Lemma 3.4.15 *All thickened clusters are finite, for $\epsilon > 0$ sufficiently small.*

Proof of Lemma 3.4.15.

By Lemma 3.4.12 there is a gap between $[-\frac{18^m}{2}, \frac{18^m}{2}]^2$ and $[-\frac{3\times 18^m}{2}, \frac{3\times 18^m}{2}]^2$, with probability at least $1/2$, by our choice of ϵ, for every m.

We now need to show that such a gap exists around the origin a.s. , for some m. If we can do this then we can conclude, by the stationarity of the circle configuration, that every point will be surrounded by some gap. We need to be careful in showing this, so that the negative information gained by the knowledge that there is no such gap, for a certain m, does not prejudice our attempts to find one in a later m. We proceed as follows.

We begin by looking whether there is a gap between the boundaries of $[-\frac{18}{2}, \frac{18}{2}]^2$ and $[-\frac{3\times 18}{2}, \frac{3\times 18}{2}]^2$. In order to do this we search for the centres of the circles that might intersect this area, in order of increasing size. Either there is no gap, in which case a (random) $K_1 < \infty$ exists, so that these boundaries are connected by circles of radius up to 18^{K_1}, or there is a gap (with probability of at least $1/2$). In the latter case we would be satisfied. In the first case we can find an M_1 so large that $[-\frac{18^{M_1}}{2}, \frac{18^{M_1}}{2}]^2$ and $[-\frac{3\times 18^{M_1}}{2}, \frac{3\times 18^{M_1}}{2}]^2$, cannot be overlapped by any circle of radius up to 18^{K_1} that could also have overlapped $[-\frac{3\times 18}{2}, \frac{3\times 18}{2}]^2$. We know nothing about larger circles.

We have no information about the circles that may connect $[-\frac{18^{M_1}}{2}, \frac{18^{M_1}}{2}]^2$ and $[-\frac{3\times 18^{M_1}}{2}, \frac{3\times 18^{M_1}}{2}]^2$; hence, the probability that there is a gap between them is again at least $1/2$. We search for a connection between the boundaries of $[-\frac{18^{M_1}}{2}, \frac{18^{M_1}}{2}]^2$ and $[-\frac{3\times 18^{M_1}}{2}, \frac{3\times 18^{M_1}}{2}]^2$, again starting by looking at the smallest circles. Either there is a gap, or there exists $K_2 < \infty$ such that there is a connection from one of these boundaries to the other using circles of size up to 18^{K_2}. In this latter case, we can find an $M_2 > M_1$ so large that $[-\frac{18^{M_2}}{2}, \frac{18^{M_2}}{2}]^2$ and $[-\frac{3\times 18^{M_2}}{2}, \frac{3\times 18^{M_2}}{2}]^2$, cannot be overlapped by any circle of radius up to 18^{K_2} that could also have overlapped $[-\frac{3\times 18^{M_2}}{2}, \frac{3\times 18^{M_2}}{2}]^2$.

We now search again for a gap, this time between $[-\frac{18^{M_2}}{2}, \frac{18^{M_2}}{2}]^2$ and $[-\frac{3\times 18^{M_2}}{2}, \frac{3\times 18^{M_2}}{2}]^2$, and repeat. At every stage we have a probability of at least $1/2$ of there being a gap, independently of the previous times. If there is no gap (which happens with a probability of at most $1/2$) then we find this out at some time and search in a larger annulus. It follows that there is almost surely a gap.

□

Finally, we can now give a proof of Theorem 3.4.7, by describing a shift invariant covering algorithm that a.s. never forms an unbounded component of covering discs, for all λ.

Proof of Theorem 3.4.7. We nearly have a covering algorithm, but everything we have done up to now depends upon λ, the density of points. As the covering should be a deterministic function of the points, we must first calculate λ in the realisation of the point configuration. We do this by setting

$$\lambda = \lim_{n\to\infty} \frac{\text{number of points in } B_n(0)}{n^2}$$

if this limit exists and is constant, which happens with probability 1, and otherwise we take $\lambda = 1$. Note that this definition is translation invariant.

Fix ϵ at half the supremum of all values of ϵ that allow all our proofs to work at this particular value of λ.

We construct *smooth* curves based upon the finite circle clusters. Consider some maximal set of circles such that, if we take the locus of points at a maximal distance of 4 from the points in the circles, then this forms a connected set, and run a disc of radius two around the outside of this set (see Figure 3.12). The disc traces out a kind of sausage shape around the clusters. We note that all such sausages must be finite by Lemma 3.4.15. In formulas we take a maximal set of circles, C, such that $\bigcup_{c\in C}\{x : |x-y| \le 4, y \in c\}$ is

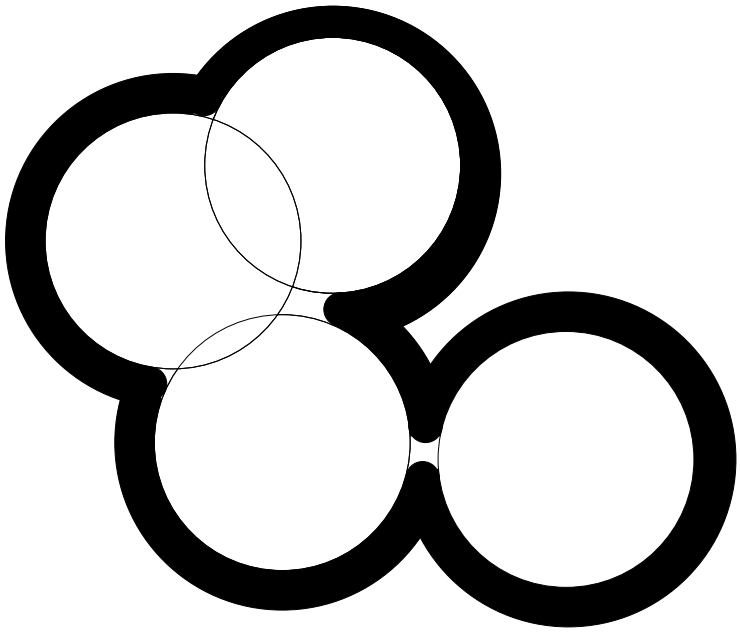

Figure 3.12: **Sausages.** By sliding a disc of radius 2 along the boundary of some cluster, we trace a kind of sausage shape.

connected, and then define the sausage to be the set on the exterior of C,

$$\{x : \exists y : |x-y| \leq 2, \exists p \in \bigcup_{c \in C} c : |y-p| = 2, \{w : |w-y| < 2\} \cap \bigcup_{c \in C} c = \emptyset\}.$$

We take the inside edge of this sausage as our curve, and note that a covering disc (of radius 1) can get arbitrarily close to any point of it without touching it.

We construct these smooth curves for each set of sufficiently close clusters, noting that they surround every region, are always finite, and never come within distance 18 of each other.

We finally cover our Poisson points as follows:

- if a point is at a distance more than 2 from every smooth curve, then we centre a covering disc at the point.

- if a point is within distance 2 of a smooth curve, then we place a disc so that its perimeter covers the point, and so that the centre of the disc is at the maximum distance away from the smooth curve. If there are a number of such possible positions, we choose the leftmost.

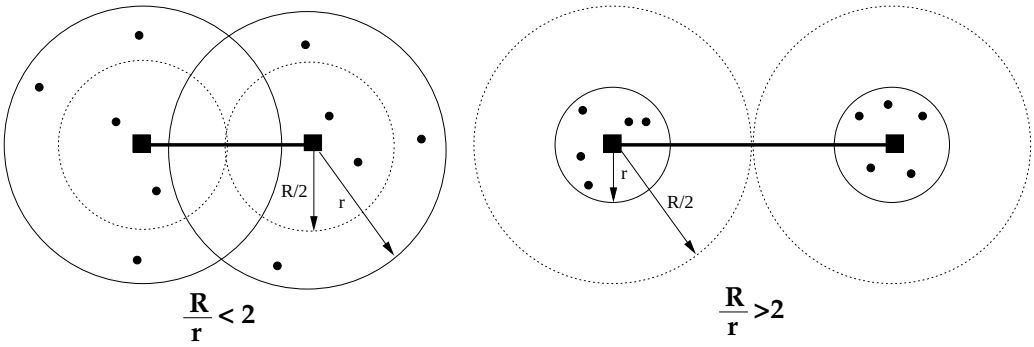

Figure 3.13: **Scaling.** Points are covered by solid line discs of radius r. Discs centres are considered connected if their distance is at most R.

It immediately follows that, for any given value of λ, a.s. there is no percolation.

□

3.5 Scaling

In this section we consider an extension that is useful to model the transmission power in wireless communication networks. We look at the percolation properties of our model, for different values of the connectivity range of the base stations and of the clients.

Let r be the clients' connectivity range and let R be the base stations' connectivity range. It follows that discs of radius r are used to cover the points of X and two disc centres are considered connected, if their distance is less than, or equal to R (see Figure 3.13). We are interested in the a.s. existence of an unbounded connected component of disc centres, for large values of the density λ of the Poisson point process. Our result is the following.

Theorem 3.5.1 (The Scaling Theorem) *Let $G \subset \mathbb{R}^2$ be the set of all vertices of a square lattice in which the distance between two neighbouring lattice vertices is δ. Call two disc centres connected if their distance is at most R. We have:*

- *CASE 1. If $\frac{R}{r} \leq 1$ then, for any $\delta > 0$, there exists a grid covering algorithm $\mathcal{A}$ that places discs only at the vertices of G, such that, for all λ, $(X, \lambda, r, \mathcal{A})$ does not percolate.*
- *CASE 2. If $1 < \frac{R}{r} < 2$ then, there exists a $\delta > 0$, depending on $\frac{R}{r}$, such that there exists a grid covering algorithm $\mathcal{A}$ that places discs only at*

the vertices of G and, for all λ, $(X, \lambda, r, \mathcal{A})$ does not percolate.

- *CASE 3. If $\frac{R}{r} = 2$ then, for any $\delta > 0$, for any grid covering algorithm $\mathcal{A}$, there exists a $\lambda_1 < \infty$, such that, for all $\lambda > \lambda_1$, $(X, \lambda, r, \mathcal{A})$ percolates.*

- *CASE 4. If $\frac{R}{r} > 2$ then, for any covering algorithm $\mathcal{A}$, there exists a $\lambda_1 < \infty$, such that, for all $\lambda > \lambda_1$, $(X, \lambda, r, \mathcal{A})$ percolates.*

Note that Case 4 of the theorem states that in a wireless network in which base stations can communicate at a distance larger than twice the maximum communication distance to the clients, an unbounded connected component forms a.s. for large values of the density of the clients, regardless of the covering algorithm used to build the cellular network.

Before giving a proof of Theorem 3.5.1, we discuss an intuitive interpretation of the theorem. Consider a fixed value of R and let r approach zero. In the limit for $r \to 0$, a covering algorithm needs to place a disc at each point of X, therefore, any covering algorithm behaves as the standard Poisson Boolean model $\{X, \lambda, R/2\}$. For this model it is known that an unbounded connected component a.s. forms, for large values of the density λ.

What Theorem 3.5.1 states is that, when r is small, the covering algorithm is constrained to place the discs almost as a Poisson point process, therefore an unbounded connected component a.s. forms, for large values of λ. On the contrary, when r is large, a covering algorithm has more freedom in placing the covering discs and percolation can be avoided.

Note that we do not need $r \approx 0$, and the covering process to behave exactly as a Poisson Boolean model, to obtain the percolation property, but as long as r is small enough that the ratio $\frac{R}{r}$ is greater than 2, Case 4 of the theorem applies, and the result on the existence of an unbounded connected component holds for any covering algorithm.

Proof of Theorem 3.5.1.
Case 1. We can restrict our attention to $\frac{R}{r} = 1$. That is because if a grid covering algorithm does not form an unbounded connected component when $\frac{R}{r} = 1$, then it does not form such a component when $\frac{R}{r} < 1$ either.

Note that, for $\frac{R}{r} = 1$, two disc centres are considered connected if and only if the corresponding discs of radius r cover each other's centres. Moreover, in order to be able to cover all points on the plane by using only grid discs of radius r, the grid spacing δ must be at most $\sqrt{2}r$.

We now consider all values of the grid spacing $\delta \leq \sqrt{2}r$, subdivided into intervals.

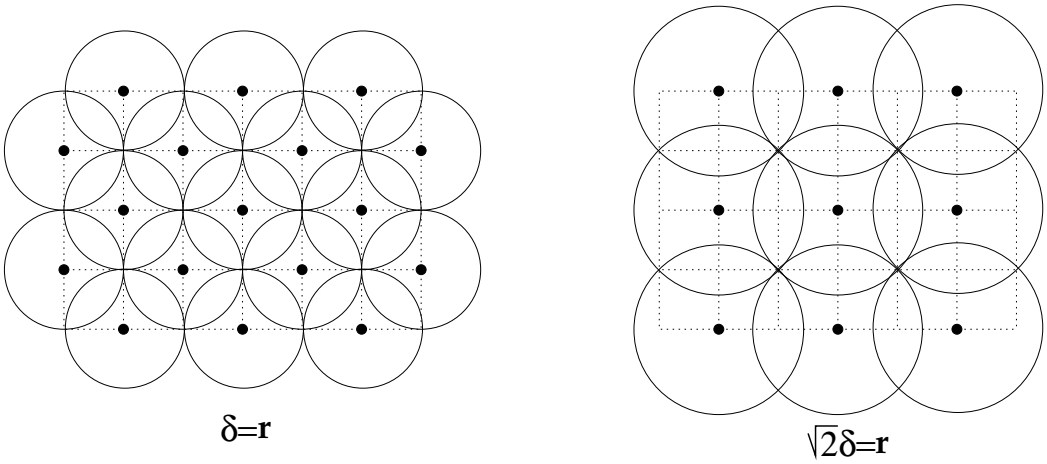

Figure 3.14: **Theorem 3.5.1, Case 1.** Two tilings of the plane by discs centred on a grid that do not cover each other's centres.

- For $r < \delta \leq \sqrt{2}r$, any grid covering algorithm places discs on the plane that do not touch each other's centres.
- For $\frac{r}{\sqrt{2}} < \delta \leq r$, consider the tiling of the plane depicted in the left part of Figure 3.14. Discs of this tiling do not cover each other's centres, therefore, any grid covering algorithm that covers all the points of X using only the grid discs depicted in the left part of Figure 3.14 does not form an unbounded connected component, a.s. , for any value of λ.
- For $r/2 < \delta \leq \frac{r}{\sqrt{2}}$, consider the tiling depicted in the right part of Figure 3.14. Discs of this tiling do not cover each other's centres, therefore, any grid covering algorithm that covers all the points of X using only the grid discs depicted in the right part of Figure 3.14 does not form an unbounded connected component, a.s., for any value of λ.

For the remaining values of δ, we can use the same tiling of the two cases depicted in Figure 3.14, scaled by the appropriate factor.

Case 2. In this case, two disc centres are considered connected if and only if the corresponding discs of radius r overlap by a region of measure at least $\epsilon > 0$, where the value of ϵ depends on the ratio $\frac{R}{r}$.

We follow a similar construction as that used to prove Proposition 3.4.3. Draw circles of radii $\{3kr, k \in \mathbb{N}\}$ around the origin, and notice that a.s. no Poisson point falls on any of these circles. Then cover the Poisson points, each with a disc of radius r, without intersecting these circles. Notice that the circles divide the plane into finite annuli, whose boundaries are not covered

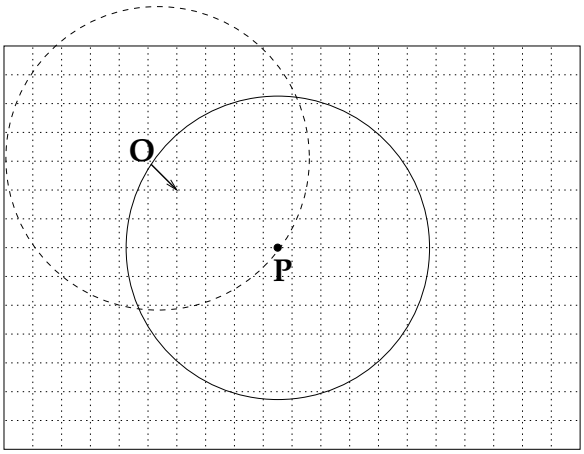

Figure 3.15: **Theorem 3.5.1, Case 2.** A Poisson point P is covered by a disc centred at point O, that is within r from P. The covering disc can be moved to a nearby grid vertex, that is inside the solid disc and is within $\sqrt{2}\delta$ from O, and still covers point P.

by discs. We now approximate this covering using a grid covering. Consider a square grid G and move each disc of the above covering to the nearest vertex of G that still allows to cover its corresponding Poisson point. Note that each disc needs to be translated by at most $\sqrt{2}\delta$. That is because a Poisson point is covered by a disc centred within r from it, and there is always a grid vertex, within radius r from the Poisson point, that is also within $\sqrt{2}\delta$ from this centre (see Figure 3.15). By this translation, some discs may intersect the boundaries of the annuli, that were previously untouched. We then take the grid size δ so small, that any two discs that intersect these boundaries do not overlap by an area of measure greater than or equal to ϵ, and are therefore not connected.

It immediately follows that, for any given value of the density λ, a.s. there is not any unbounded connected component for this covering.

Case 3. This case is proven by Theorem 3.4.3.

Case 4. In this case, two discs are considered connected if and only they are at a distance of at most $R - 2r$ (see right hand side of Figure 3.13).

We construct a mapping from the covering discs to a discrete site percolation model. Consider a partition of the plane into boxes of side length $\frac{\epsilon}{\sqrt{2}}$, with $0 < \epsilon \leq \frac{R}{2} - r$. Note that if some point of X falls inside a box of the partition, then it must be covered by a disc of radius r, and therefore the entire box is covered by a disc of radius $\frac{R}{2}$ (see Figure 3.16).

Consider now each $\epsilon \times \epsilon$ box as a site of a site percolation model. Call

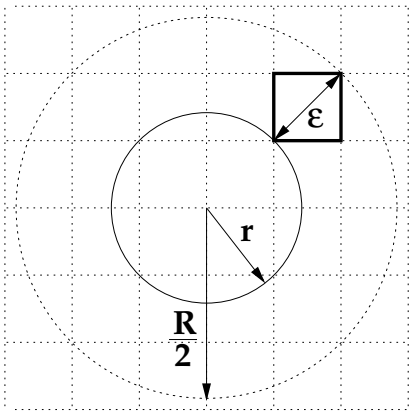

Figure 3.16: **CASE 4.** If some point of X falls inside a box of side length $\frac{\epsilon}{\sqrt{2}}$, then the entire box is covered by a disc of radius $\frac{R}{2}$.

the site occupied, if there is at least a point of X situated inside the box. Clearly, the occupancy of a site is independent of the occupancy of other sites, and, if two adjacent sites are both occupied, then the corresponding covering discs form a connected component. Next, we choose λ large enough that the probability of a site being occupied is larger than p_c, where p_c is the critical probability for site percolation on a square lattice. The a.s. existence of an unbounded connected component of covering discs immediately follows. □

3.6 Optimal algorithms

In this section we explore the notion of an optimal algorithm, i.e. one which uses as few discs as possible. We first consider n-square algorithms, and show that they are nearly optimal, then we describe an algorithm which really is optimal. Note that n may take any positive real value. The work in this section is philosophically close to that in Yukich (1998), for example.

Fix the density of points, λ, and r, and extend the definition of the density of a covering, $\mathcal{A}$, to be

$$\delta_{\mathcal{A}} = \lim_{n\to\infty} \frac{\text{number of discs centred in } B_n}{n^2}$$

if this exists and is a constant a.s., and ∞ otherwise. We then define the *optimal density* to be

$$\delta^{opt} = \inf_{\mathcal{A}} \delta_{\mathcal{A}},$$

where the infimum runs over all covering algorithms. We are interested in whether there exists an *optimal algorithm*, $\mathcal{A}^{opt}$, for which, a.s. , $\delta^{opt} = \delta_{\mathcal{A}^{opt}}$. We would not expect such an algorithm to have a finite horizon.

First we show that n-square algorithms can get as close as we like to the optimal density. Define δ_n to be the density of discs under an n-square algorithm. By ergodicity $\delta_n < \infty$ exists. Notice that δ_n does not depend upon the particular n-square algorithm we choose.

Theorem 3.6.1 *Given $\epsilon > 0$, there exists n_ϵ such that $\delta_{n_\epsilon} < \delta^{opt} + \epsilon$, and hence $\delta^{opt} = \inf_n \delta_n$.*

Proof of Theorem 3.6.1 We prove this theorem by contradiction, so suppose that we can find an ϵ such that there is no n-square covering with density between δ^{opt} and $\delta^{opt} + \epsilon$. We can find another covering, $\hat{\mathcal{A}}$ say, with density $\delta_{\hat{\mathcal{A}}} \in [\delta^{opt}, \delta^{opt} + \epsilon/4]$, by the definition of δ^{opt}.

Choose $\gamma > 0$ such that $(1-\gamma)(\delta^{opt}+\epsilon/2)+\gamma(1+\epsilon/2)/r^2 < \delta^{opt}+\epsilon$. Note that the number of discs necessary to cover an n-square is at most $\lceil n/r \rceil^2$.

As $\delta_{\hat{\mathcal{A}}} = \lim_{n\to\infty} \frac{\text{number of discs centred in } B_n}{n^2}$, a.s, we can choose $n_\epsilon > 1$ sufficiently large that a) the number of discs centred in B_{n_ϵ} is less than $n_\epsilon^2(\delta^{opt} + \epsilon/2)$ with probability larger than $1 - \gamma$, and b) $\lceil n/r \rceil^2 (1/n)^2 \leq (1 + \epsilon/2)/r^2$.

Given a finite set of points there are a finite number of distinct possible coverings of those points, where we call two coverings *distinct* if there exist a set of points covered by one disc in one of the coverings but covered by two or more in the other. Coverings that are not distinct are *equivalent*. We will occasionally work with the equivalence classes of theses coverings.

We next define a covering of the box B_{n_ϵ} which is based upon $\hat{\mathcal{A}}$ but is independent of the points outside B_{n_ϵ}. Given a point configuration, π, in B_{n_ϵ} there is a finite set of equivalent classes of coverings of these points. Let S_π be the subset of equivalence classes which occur with positive probability if we use $\hat{\mathcal{A}}$ to cover $\pi \cup X|'_{B^c_{n_\epsilon}}$, where $X|'_{B^c_{n_\epsilon}}$ is a Poisson process on $B^c_{n_\epsilon}$ independent of π and X. In each equivalence class all coverings use the same number of discs so we can choose an equivalence class from S_π in which the number of discs used is minimal, according to some deterministic rule. Choose a covering from this class, according to some other deterministic rule. This is the covering we use to cover B_{n_ϵ}. Note that this is independent of the actual point process outside B_{n_ϵ}. The expected number of discs required to cover the points in B_{n_ϵ} under this algorithm can be at most that under $\hat{\mathcal{A}}$.

We divide up the plane into squares of size n_ϵ, and cover each square independently using the same algorithm in each as we use on B_{n_ϵ}. For those squares for which this requires at most $n_\epsilon^2(\delta^{opt} + \epsilon/2)$ discs, we use this cov-

ering. In the other squares we cover optimally, which means that we have a density of at most $(1+\epsilon/2)/r^2$ on these squares.

We have created an algorithm that covers each square of size n_ϵ independently, and which therefore cannot have a density less than δ_{n_ϵ}. However, the density of the covering is at most $(1-\gamma)(\delta^{opt}+\epsilon/2)+\gamma(1+\epsilon/2)/r^2 < \delta^{opt}+\epsilon$, and we have a contradiction. □

Next, we prove a proposition that extends the previous theorem and will be useful later.

Proposition 3.6.2 *Let δ_n be the density of discs under an n-square algorithm. Then*

$$\lim_{n\to\infty} \delta_n = \delta^{opt} \ .$$

Proof of Proposition 3.6.2

We know that $\delta^{opt} = \inf_n \delta_n$, and that $\delta_n n^2$ is the expected number of discs needed to cover an n-square. For the sake of contradiction suppose that there exist $\epsilon > 0$ and a sequence $\{t_1, t_2, \ldots\}, \lim_i t_i = \infty$ such that $\delta_{t_i} > \delta^{opt} + \epsilon$ for all i. However we can choose α so that $\delta_\alpha < \delta^{opt} + \epsilon/3$. We can also choose i so large that $\left(1 - \left\lfloor \frac{t_i}{\alpha} \right\rfloor^2 \left(\frac{\alpha}{t_i}\right)^2\right)\lambda < \epsilon/3$. The reason we need this will become clear shortly.

We cover the square B_{t_i} as follows. We first divide as much of the square as possible into squares of size α. Each of these we cover optimally. We have an area of $(t_i^2 - \alpha^2 \left\lfloor \frac{t_i}{\alpha} \right\rfloor^2)$ left, and each of the points in this area we cover with one disc. This gives us a covering with expected density,

$$\left\lfloor \frac{t_i}{\alpha} \right\rfloor^2 \left(\frac{\alpha}{t_i}\right)^2 \delta_\alpha + \left(1 - \left\lfloor \frac{t_i}{\alpha} \right\rfloor^2 \left(\frac{\alpha}{t_i}\right)^2\right)\lambda \leq \delta^{opt} + \epsilon/3 + \epsilon/3.$$

However, the minimal expected density for any algorithm covering the box B_{t_i}, $\delta_{t_i} > \delta^{opt} + \epsilon$, so we have a contradiction. □

Note that it is still not clear a priori that an optimal algorithm should exist. The existence of an optimal density, defined as the infimum over all attainable densities, does not have to be attainable itself. However, we have the following theorem:

Theorem 3.6.3 *There exists an optimal algorithm.*

We give first the algorithm that we claim is optimal, and then prove that this is so in a number of steps.

The algorithm

The basic idea is to recursively find coverings of points in boxes, which are part of optimal coverings of points in larger boxes, incrementally covering the whole plane. Consider the boxes $B_n, n \in \mathbb{N}$. In each box there is a finite number of distinct possible optimal coverings of the points in that box, and from now on we work with the equivalence classes of these coverings.

Let A_n be an optimal covering of B_n chosen according to some rule, and let $A_n|_m$ be the covering of the points of B_m induced by A_n, by which we mean the covering consisting of all discs of A_n that cover at least one point of B_m. Take B_1 and consider the sequence of coverings $A_n|_1, n \in \mathbb{N}$. There are only finitely many equivalent coverings of B_1, and at least one of these must appear infinitely many times in $A_n|_1, n \in \mathbb{N}$. Choose such an equivalence class and cover B_1 using a covering from this class. Let $I_1 \subset \mathbb{N}$ be the infinite set of indices, such that $A_n|_1, n \in I_1$ is in the chosen equivalence class. Let J_1 be the smallest element of this set that is larger than 1.

We can repeat this exercise to find a covering of B_{J_1}. We consider the sequence of equivalence classes of coverings $A_n|_{J_1}, n \in I_1$. Again this set is finite, so we can find an infinite set $I_2 \subseteq I_1$, such that all the coverings $A_n|_{J_1}, n \in I_2$ are in the same equivalence class. We fix a covering for B_{J_1} from this class, which is necessarily consistent with the covering we have already chosen for B_1. Let J_2 be the smallest element of I_2 that is larger than J_1.

We repeat this procedure in the natural way. Every time we fix a covering of a box corresponding to the lowest element of some index set, and take a new index set which is an infinite subset of the previous one. We find in this way a covering of the whole plane.

It is not obvious that this covering should be optimal. We have a sequence of boxes, $\{B_1, B_{J_1}, B_{J_2}, \ldots\}$ which contain coverings that are parts of optimal coverings of larger boxes, but there is no a priori guarantee that the limit $\lim_{i\to\infty} \frac{\text{number of discs in } B_i}{i^2}$ will be δ^{opt}. To show that it really is optimal we need a few definitions and a lemma.

For a bounded subset of the plane, V, let $\tilde{V}$ be the set $\{x : \exists y \in V \text{ such that } |x - y| \leq 2r\}$. Let N_V be the (random) number of discs in an optimal covering of the points in V. If $W \supseteq \tilde{V}$ is another such subset then, given an optimal covering algorithm of W, let $N_W|_V$ be the number of discs in the covering of W that also cover points in V.

Lemma 3.6.4 *Let A and B be bounded subsets of the plane such that $B \supseteq \tilde{A}$. Fix optimal covering algorithms for $\tilde{A}$ and B. Then for every point configuration $N_B|_A \leq N_{\tilde{A}}$.*

Proof of Lemma 3.6.4 Notice first that we can find a covering of B by

covering all the points in $\tilde{A}$ optimally and then use an optimal covering of B to cover the points in $B/\tilde{A}$. This gives us:

$$N_B \leq N_{\tilde{A}} + N_B|_{B/\tilde{A}}.$$

Secondly, notice that any point in $B/\tilde{A}$ is at least $2r$ from any point in A. This implies that each disc used in an optimal covering of B can cover points in only one of these sets. Thus,

$$N_B|_A + N_B|_{B/\tilde{A}} \leq N_B.$$

Combining these inequalities gives us the result. □

Let us take a box, B_t, as A in the lemma, and B_{J_i} as B, where $B_{t+2r} \subset B_{J_i}$. Then the lemma says that the number of discs covering points in B_t under an optimal covering of B_{J_i} is at most the number of discs covering points in B_{t+2r} under an optimal covering of this area. We note that this last quantity is at most the optimal number of discs covering B_t plus the number of points in B_{t+2r}/B_t (call this number P_t). Hence,

$$N_{B_{J_i}}|_{Bt} \leq N_{B_{t+2r}} \leq N_{B_t} + P_t.$$

Let M_t be the number of discs in the covering that we claim is optimal that cover points in B_t. Note that $M_t = N_{B_j}|_{B_t}$ for all $j > t + 2r$ with $j \in \{J_1, J_2, \ldots\}$. Therefore we can write:

$$\frac{N_{B_t}}{t^2} \leq \frac{M_t}{t^2} \leq \frac{N_{B_t} + P_t}{t^2}.$$

Notice that $\frac{P_t}{t^2}$ goes to 0 as t goes a.s. to infinity, as follows from, for example, Chebyshev's inequality.

It follows that to prove Theorem 3.6.3 it is enough to show:

Lemma 3.6.5

$$\lim_{t\to\infty} \frac{N_{B_t}}{t^2} = \delta^{opt}, \ \ a.s.$$

Proof of Lemma 3.6.5 We use a continuous sub-additive ergodic theorem. See Akcoglu and Krengel (1981), in particular Theorem 2.8, for more details.

Let N_R be the minimum number of discs required to cover all the points in the rectangle R, and suppose we have two disjoint rectangles, R_1 and R_2. Then we note that:

$$N_{R_1 \cup R_2} \leq N_{R_1} + N_{R_2}.$$

It then follows that

$$\lim_{t\to\infty} \frac{N_{B_t}}{t^2}$$

exists and is a constant, a.s. Proposition 3.6.2 together with the fact that $N_{B_t}/t^2, t > t'$ is bounded above by $\lceil \frac{t}{r} \rceil^2 \frac{1}{t^2}$, for any $t' > 0$, gives by dominated convergence that this limit is δ. □

3.7 Open problems

We would like to mention a number of open problems:

- For which classes of algorithms does there exist a critical density? By this we mean a a critical value λ_c, such that percolation occurs for $\lambda > \lambda_c$ and does not occur for $\lambda < \lambda_c$.
- For which classes of algorithms is the infinite cluster unique? In other words, when do we have either 0 or 1 infinite cluster, a.s.?
- We have shown in Theorem 3.4.5 that if we have an algorithm with a finite horizon, which is shift invariant under two linearly independent shifts and has a bounded density of discs, then we must have percolation for λ high enough. We have also shown in Theorem 3.4.7 that we can have a completely shift invariant covering algorithm with an unbounded density of discs and no finite horizon that does not percolate, even for high values of λ. Do we have percolation for λ high enough for a finite horizon algorithm invariant under a pair of shifts, with an unbounded density of discs? Do we necessarily have percolation for λ high enough if we have a bounded density of discs and shift invariance but no finite horizon?

Acknowledgements. The authors would like to thank Richard Gill, who suggested the intuitive interpretation of Theorem 3.5.1, reported in Section 5. We would also like to thank Johan Segers for reading an early version of this article very carefully, and the many interesting comments and observations that resulted. Massimo Franceschetti would like to thank his fellow student Matthew Cook, for many interesting discussions, and for reading and providing feedback on some of the proofs. We would also like to thank the referee of this paper for scrutinising this paper scrupulously.

Bibliography

[1] P.J. Agarwal and M. Sharir (1998). Efficient Algorithms for Geometric Optimization. *ACM Computing Surveys,* 30(4), 412-458.

[2] M.A. Akcoglu and U. Krengel (1981). Ergodic Theorems for superadditive processes. *Journal für die Reine und Angewandte Mathematik,* 323, 53-67.

[3] R. Durrett and D. Griffeath (1983). Supercritical contact processes on $\mathbb{Z}$. *Annals of Probability,* 11(1), 1-15.

[4] P. Erdős and A. Rényi (1959). On Random Graphs I. *Publicationes Mathematicae Universitatis Debreceniensis,* 6, 290-297.

[5] P. Erdős and A. Rényi (1960). On the Evolution of Random Graphs. *Publications of the Mathematical institute of the Hungarian Academy of Sciences,* 5,17-61.

[6] P. Erdős and A. Rényi (1961). On the Evolution of Random Graphs. *Bulletin of the International Institute of Statistics of Tokio,* 38, 343-347.

[7] P. Erdős and A. Rényi (1961). On the Strength of Connectedness of a Random Graph. *Acta Mathematica Acadamiae Scientiarum Hungaricae,* 12,261-267.

[8] M. Franceschetti, M. Cook and J. Bruck (2001). A Geometric Theorem for Approximate Disk Covering Algorithms. Submitted for publication. Available online at: `http://www.paradise.caltech.edu/ETR.html`.

[9] M. Gerla and J.T. Tsai (1995). Multicluster, mobile, multimedia radio network. *ACM/Baltzer Wireless Networks,* 1(3), 255-265.

[10] E. N. Gilbert (1961). Random Plane Networks. *Journal of the Society for Industrial and Applied Mathematics,* 9(4), 533-543.

[11] P. Gupta and P.R. Kumar (1998). Critical Power for Asymptotic Connectivity in Wireless Networks. *Stochastic Analysis, Control, Optimization and Applications: a Volume in Honor of W.H. Fleming. W.M. McEneany, G. Yin, and Q. Zhang editors, Birkhäuser, Boston.*

[12] P. Gupta and P.R. Kumar (2000). The Capacity of Wireless Networks. *IEEE Transactions on Information Theory,* 46(2), 388-404.

[13] P. Hall (1985). On Continuum Percolation. *The Annals of Probability,* 13(4), 1250-1266.

[14] P. Hall (1988). Introduction to the Theory of Coverage Processes. *J. Wiley and sons, New York.*

[15] M. Penrose (1997). The longest edge of the random minimal spanning tree. *The Annals of Applied Probability,* 7(2), 340-361.

[16] H. Kesten (1980). Percolation Theory for Mathematicians, *Birkhäuser.*

[17] R. Meester and R. Roy (1994). Uniqueness of Unbounded Occupied and Vacant Components in Boolean Models. *The Annals of Applied Probability,* 4(3), 933-951.

[18] R. Meester and R. Roy (1996). Continuum Percolation. *Cambridge University Press.*

[19] M. V. Menshikov (1986). Coincidence of Critical Points in Percolation Problems. *Soviet Mathematics Doklady,* 24, 856-859.

[20] R. Roy (1990). The Russo-Seymour-Welsh Theorem and the Equality of the Critical Densities and Critical Dual Densities for Continuum Percolation on $\mathbb{R}^2$. *The Annals of Probability,* 18, 1563-1575.

[21] J. Segers. Personal Communication

[22] J.E. Yukich (1998). Probability Theory of Classical Euclidean Optimization Problems. *Lecture Notes in Mathematics 1675, Springer.*

Chapter 4

On the construction of $\mathbb{Z}^{\mathbb{Z}^d}$-valued systems with unbounded transition rates

Lorna Booth and Corrie Quant

We construct certain particle systems on $\mathbb{Z}^{\mathbb{Z}^d}$. These systems start in configurations given by a stationary, ergodic law, such that the expectation of the absolute number of particles per site is finite. Particles are added and removed at rates which are functions of the configuration of particles. These functions are of bounded variation and are such that the rate at which particles are added at each is site is almost surely finite. Negative numbers of particles are allowed.

4.1 Introduction

Interacting particle systems are often used to model aspects of the physical world, but mathematicians and physicists have different approaches to these systems. Sometimes physicists discuss the invariant measure of a certain system, while mathematicians are worrying whether the system even exists, and trying to construct a model fitting the informal description of the physicists.

The construction of systems with bounded local transition rates has been widely studied, in for example Liggett [4, 5]. Examples of constructions of particle systems with unbounded and non-local transition rates can be found

in Liggett [3], Andjel [1], Maes et al. [6] and Meester and Quant [7]. These constructions use the monotonicity of the systems.

In this paper we consider the existence of certain infinite particle systems on $\mathbb{Z}$, although our results can easily be extended to systems on $\mathbb{Z}^d$. The precise result can be found in Section 4.2. We allow ourselves to have negative, as well as positive, numbers of particles. The systems we construct are not necessarily nearest-neighbour systems or monotone systems and may have unbounded rates. We will give conditions under which such a system with a given formal generator (which comes down to a given informal description of the transition rates) can be constructed. Negative particles allow us to model an interface, where the height differences are interpreted as particles. In Section 4.4 we give an example of the construction of a system that we can handle in this way (under certain restrictions): the bricklayer model of Balàzs [2].

Informally described, we let our systems start from an initial configuration chosen according to an ergodic stationary measure, for which the expectation of the absolute value of the number of particles per site is finite. Particles are added to or removed from a configuration with a rate that is a function of bounded variation of the configuration of particles already present. This is done in a spatially stationary way. The rate at which particles are added and removed at each site will then be finite, a.s.

In Section 4.3 we consider the class of rate functions we allow, Section 4.4 contains the bricklayer example, and finally Section 4.5 contains the proof of our result.

4.2 Notation and results

We write Ω for the subset of configurations of $\mathbb{Z}^{\mathbb{Z}}$ for which

$$\limsup_{n\to\infty} \frac{|\omega(-n)| + \cdots + |\omega(n)|}{2n+1} < \infty.$$

The set Ω will be the state space of our system, which we equip with the product topology and Borel σ-algebra. Given $\omega \in \Omega$, we denote its lth component by $\omega(l)$. The shift $T : \Omega \to \Omega$ is defined by $T\omega(l) = \omega(l-1)$, for all $l \in \mathbb{Z}$, and, logically enough, $T^j\omega(l) = \omega(l-j)$, for $j, l \in \mathbb{Z}$. Whenever we talk about stationarity, it will be with respect to the shift. Define $|z|$ to be the sum of the absolute values of the components of $z \in \mathbb{R}^{\mathbb{Z}}$. Let $||z||(l) = |z(l)|$, for all $l \in \mathbb{Z}$, giving $||z||$ components that are the absolute values of those of z. For a constant c, $c^+ = c\mathbb{1}_{\{c\geq 0\}}$ and $c^- = -c\mathbb{1}_{\{c<0\}}$, where $\mathbb{1}_A$ is the indicator of the event A. Similarly the positive and negative parts of

$i \in \mathbb{Z}^{\mathbb{Z}}$, i^+ and i^-, are defined by $i^+(l) = i(l)^+$ and $i^-(l) = i(l)^-$ respectively. Given $a, b \in \mathbb{Z}^{\mathbb{Z}}$, we write the inner product of a and b as $a \cdot b = \sum_{l \in \mathbb{Z}} a(l)b(l)$, when this is well defined. Let $\mathbf{1}_l \in \mathbb{Z}^{\mathbb{Z}}$ be given by $\mathbf{1}_l(l) = 1$ and $\mathbf{1}_l(q) = 0$, for $q \neq l$. If we say that $x \leq y$ for $x, y \in \mathbb{Z}^{\mathbb{Z}}$, we mean that $x(l) \leq y(l)$, for all l. A process with arrival rates that are monotone increasing (in the configurations) we call a *monotone process*. This is slightly stronger than the standard definition of monotonicity (also known as attractivity) in the literature.

Let ν be an ergodic stationary measure on Ω with

$$\mathbb{E}_\nu(|\omega(0)|) := \int_\Omega |\omega(0)| \, d\nu(\omega) < \infty,$$

and suppose that I is a countable subset of $\mathbb{Z}^{\mathbb{Z}}$, with the property that the elements of I have only finitely many non-zero coordinates. For each $i \in I$ we have an associated *rate function* $f_i : \Omega \to \mathbb{R}_{\geq 0}$, satisfying certain conditions, which we specify later. We start with an initial configuration $\omega \in \Omega$, chosen according to ν. Informally the dynamics of the system can be described as follows. For all $j \in \mathbb{Z}$, $i \in I$ we add $i \in I$ at position j (that is we add $T^j i$), with rate $f_i(T^{-j}\omega)$ to give $w + T^j i$. We do not require that $i \in I$ are uniformly 0 outside a certain interval, that I is finite, or that f_i depends only upon finitely many coordinates. We write ν^* for the law of the process when the initial configuration is chosen according to ν.

The description above is informal in the sense that it it does not a priori give rise to a well defined particle system on Ω. The rates need not be bounded which might cause problems. However we will construct such a system, and make precise what we mean by this in Theorem 4.2.1.

To get some feeling for the systems we can construct, we give an example. We could take:

$$i_1(l) = \begin{cases} -1 & l = 3 \\ 4 & l = 5 \\ 0 & \text{otherwise} \end{cases}$$

and

$$i_2(l) = \begin{cases} -2 & l = 1 \\ -2 & l = 2 \\ 0 & \text{otherwise,} \end{cases}$$

with rates functions given by

$$f_{i_1}(\omega) = \sum_{j \in \mathbb{Z}} \frac{\omega(j)}{j^2 + 4}$$

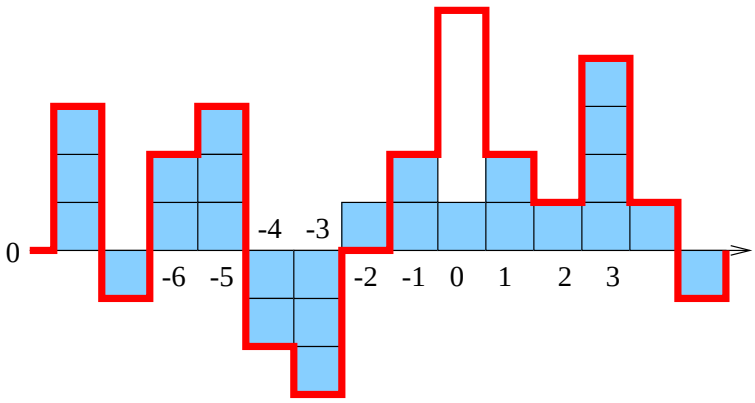

Figure 4.1: Adding i_1 at -5: the original configuration is shaded, while the new configuration is given by the thick grey line

and

$$f_{i_2}(\omega) = 3\sqrt{|\omega(2) + \omega(15)|}.$$

The effect of the addition of i_1 at -5 is shown in Figure 4.1. Another example, the bricklayer model, is given in Section 4.4.

The class of functions from which we will choose our rate functions we call $\mathcal{F}$, and we will specify it exactly in the next section. The sub-class of $\mathcal{F}$ possibly of most use are the functions of bounded variation. A function $f(\omega)$ is of *bounded variation* if there exists $k \in \mathbb{R}^{\mathbb{Z}}_{\geq 0}$, $|k| < \infty$ such that,

$$|f(\omega + \mathbf{1}_j) - f(\omega)| \leq k(j),$$

for all configurations $\omega \in \Omega$ and $j \in \mathbb{Z}$. A collection of such functions $\{f_i : i \in I\}$ is allowed if they fulfill certain conditions, which will imply that expected total rate at which particles at a site are added and removed is finite, when we start with a finite expected absolute number of particles per site and a stationary, ergodic configuration. We will make this precise in Section 4.3.

We can now state the main theorem:

Theorem 4.2.1 *Let ν be an ergodic stationary measure on Ω with $\mathbb{E}_\nu(|\omega(0)|) < \infty$ and let $h : \Omega \to \mathbb{R}$ and $z \in \mathbb{R}^{\mathbb{Z}}_{\geq 0}$, $|z| < \infty$, be such that for all $j \in \mathbb{Z}$, $\omega \in \Omega$,*

$$|h(\omega + \mathbf{1}_j) - h(\omega)| \leq z(j).$$

Let $I \subset \mathbb{Z}^{\mathbb{Z}}$ be countable and let $\mathcal{F}_I := \{f_i : i \in I\}$ be nice, as defined in Section 4.3. Let the formal generator G for functions $g : \Omega \to \mathbb{R}$ and $\omega \in \Omega$

be defined by

$$Gg(\omega) = \sum_{i \in I,\ j \in \mathbb{Z}} f_i(T^{-j}\omega)(g(\omega + T^j i) - g(\omega)).$$

Then, ν-almost surely, there exists a Markov process $(\omega_t)_{t \geq 0}$ on Ω with semigroup $S(t)$ defined by

$$S(t)h(\omega) := \mathbb{E}(h(\omega_t)).$$

Moreover,

$$\nu^* \left(\lim_{m \downarrow 0} \frac{S(m)h(\omega_t) - h(\omega_t)}{m} = Gh(\omega_t) \text{ for all } t \right) = 1.$$

We prove this theorem in Section 4.5, using the following idea. We call the Markov process $(\omega_t)_{t \geq 0}$ the *original system*. We will first construct a more complicated system than the original system, which we call the *political system*. It has two types of particles, red and blue, but is easier to construct as it will have monotonicity in its rates, and it will only have positive particles. We will show that, under suitable conditions, the number of particles at each site in this system is finite, using a comparison with a spatial branching process. Our original system will be defined as the difference between the number of red and the number of blue particles in the political system, which is then well defined.

The relationship between the semigroup and generator can then be shown by coupling this political process to another, simpler system. This part of the proof is quite technical.

4.3 The class $\mathcal{F}$

Let Ω^+ be the subset of $\mathbb{N}^{\mathbb{Z}}$ consisting of elements x for which

$$\limsup_{n \to \infty} \frac{x(-n) + \cdots x(n)}{2n+1} < \infty.$$

A function $f : \Omega \to \mathbb{R}_{\geq 0}$ is in $\mathcal{F}$ if there exist a pair of functions $f^r, f^b : (\Omega^+ \times \Omega^+) \to \mathbb{R}_{\geq 0}$ and a constant $c \in \mathbb{R}$ such that

- For all $x, y \in \Omega^+$, $f(x - y) = c + f^r(x, y) - f^b(x, y)$,
- $f^r(x, y)$ and $f^b(x, y)$ are monotone increasing in every coordinate for $x, y \in \Omega^+$,

- there exists $k \in \mathbb{N}^{\mathbb{Z}}$, with $|k| < \infty$, such that $f^r(x,y) \leq k \cdot (x+y)$ and $f^b(x,y) \leq k \cdot (x+y)$, for all x, $y \in \Omega^+$.

When we write $f_i(x-y) = c_i + f_i^r(x,y) - f_i^b(x,y)$ for $f_i \in \mathcal{F}$, $i \in I$, the associated k is denoted by k_i.

These conditions might seem strange, but recall that we will define the original process as the difference of two monotone processes of red and blue particles. The functions $c_i^+ + f_i^r(x,y)$, $i \in I$, will be the growth rates of the red particles and $c_i^- + f_i^b(x,y)$, $i \in I$, of the blue. If we let the changes in the red and blue processes occur at the same times (when possible), the changes in the original process will occur according to the "correct" rates.

We say that a collection of rate functions $\mathcal{F}_I = \{f_i : i \in I\}$ is *nice* if

- $\mathcal{F}_I \subset \mathcal{F}$,
- for each function f_i a k_i can be found as above, and $\kappa = \sum_{i \in I} |k_i||i| < \infty$,
- $\gamma = 2\sum_{i \in I} |c_i||i| < \infty$, with the constants $c_i, i \in I$ as above.

The last two conditions will imply that the expected rate of addition and removal of particles at any site is finite under the appropriate starting conditions.

Which functions are in $\mathcal{F}$? The definition of the class given above may seem complicated, but is not as restrictive as it might seem. We explore the class a little, starting with an example.

Take the function $f_i(\omega) = \sqrt{\omega(0)\mathbb{1}_{\{\omega(0)>0\}}}$. We can write

$$\begin{aligned} f_i(x-y) &= \sqrt{(x(0)-y(0))\mathbb{1}_{\{x(0)-y(0)>0\}}} \\ &= (\sqrt{(x(0)-y(0))\mathbb{1}_{\{x(0)-y(0)>0\}}} + \mathbf{1}_0.y) - (\mathbf{1}_0.y) \end{aligned}$$

for $(x,y) \in \Omega^+ \times \Omega^+$. Note that the terms in the last expression in brackets are monotone in x and y, and less than $\mathbf{1}_0 \cdot (x+y)$. This tells us that this function is in $\mathcal{F}$.

In this example we were able to express the function of interest as the difference of monotone increasing functions by both adding and subtracting a linear function of x and y. This technique can be extended to functions of bounded variation, in the following way.

A function of bounded variation satisfying $|f(\omega + \mathbf{1}_j) - f(\omega)| \leq k^*(j)$, for all $\omega \in \Omega$, $j \in \mathbb{Z}$ and some $k^* \in \mathbb{R}_{\geq 0}^{\mathbb{Z}}$ such that $|k^*| < \infty$, we can write as

$$f(\omega) = f^1(\omega) - f^2(\omega) + c,$$

where $f^1(\omega) = k^* \cdot \omega$ and $f^2(\omega) = k^* \cdot \omega - f(\omega)$ are monotone increasing, $f^1(\mathbf{0}) = f^2(\mathbf{0}) = 0$ and such that $|f^1(\omega + \mathbf{1}_j) - f^1(\omega)| = k^*(j)$ and $|f^2(\omega + \mathbf{1}_j) - f^2(\omega)| \leq 2k^*(j)$.

If we set $k(j) = 2k^*(j)$ we can then write

$$f^r(x,y) = (f^1(x-y)\mathbb{1}_{f^1(x-y)\geq 0} + k \cdot y) + (-f^2(x-y)\mathbb{1}_{f^2(x-y)\leq 0} + k \cdot x) + c^+$$
$$f^b(x,y) = (-f^1(x-y)\mathbb{1}_{f^1(x-y)\leq 0} + k \cdot x) + (f^2(x-y)\mathbb{1}_{f^2(x-y)\geq 0} + k \cdot y) + c^-$$

Note that each of the terms in brackets is monotone increasing in x and y and that $f^r(x,y) - f^b(x,y) = f(x-y)$.

4.4 The bricklayers model - an example

In Balàzs [2] a *bricklayer model* is introduced, which is a process on $\mathbb{Z}^{\mathbb{Z}}$. The system is a nearest neighbour system with unbounded rates. Balàzs does not construct this system, but (under the assumption it exists) he achieves various results on the invariant measure of the process, which turns out to be a product measure. These include interesting shock solutions. We can construct the system if the rate functions do not grow too fast.

Configurations of the process are denoted by $\omega \in \Omega$ and the random state of the process at time t and with initial configuration ω is denoted by ω_t. We can construct the model in the case that the rates are of bounded variation, and the initial configuration is chosen according to an ergodic stationary measure such that the expectation of $|\omega(0)|$ is finite (in this case the process will a.s. concentrate on Ω).

We begin by giving an informal description of the bricklayer process. Let $r : \mathbb{Z} \to \mathbb{R}_{\geq 0}$ be a monotone increasing function that is of bounded variation, i.e. there should exist a $\kappa < \infty$ such that $|r(z+1) - r(z)| \leq \kappa$, for all $z \in \mathbb{Z}$, and that has the property that for all $l \in \mathbb{Z}$,

$$r(l)r(-l+1) = 1.$$

The dynamics of the process can be described as

$$(\omega(l), \omega(l+1)) \to (\omega(l) - 1, \omega(l+1) + 1) \text{ at rate } r(\omega(l)) + r(-\omega(l+1)).$$

The process can be interpreted as a bricklayer process in the following sense. Think of an infinite wall, built of bricks, as in Figure 4.2. We identify the surface of the wall with $\omega \in \Omega$, where $\omega(l)$ denotes the height difference between the column of bricks between the sites $(l-1)$ and l and the sites l and $(l+1)$.

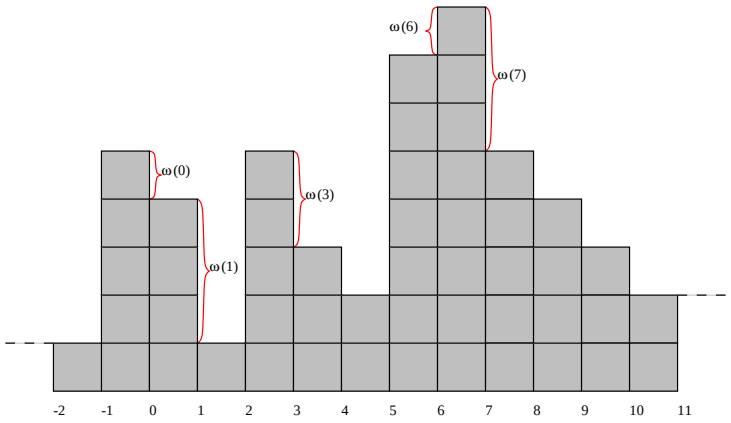

Figure 4.2: Bricklayer process

Imagine that you take a walk along the top of this wall, from right to left. Then the height difference is positive if you go up, and negative if you go down. For example, in Figure 4.2, $\omega(0) = 1$, $\omega(1) = 3$, $\omega(3) = 2$, $\omega(6) = -1$ and $\omega(7) = 3$. At each site l, a bricklayer is present, putting a brick to his right at rate $r(\omega(l))$ and to his left at rate $r(-\omega(l))$.

Define i by

$$i(j) = \begin{cases} -1 & \text{if } x = 0, \\ +1 & \text{if } x = 1, \\ 0 & \text{if } x \neq 0, 1, \end{cases}$$

the function f_i by

$$f_i(\omega) = r(\omega(0)) + r(-\omega(1)),$$

and write $I = \{i\}$. Note that $|i| = 2 < \infty$. Define $k_i = \mathbf{1}_0 + \mathbf{1}_1$ and note that $f_i(\omega) \leq k_i \cdot ||\omega||$. This means that $|k_i||i| = 2 \times 2 < \infty$, so our function is in $\mathcal{F}$ and $\mathcal{F}_I$ is nice. Then the formal generator $\tilde{G}$ of the bricklayer process is, for $g : \Omega \to \mathbb{R}$, given by

$$\tilde{G}g(\omega) = \sum_{j \in \mathbb{Z}} f_i(T^{-j}\omega)(g(\omega + T^j i) - g(\omega)).$$

Since the rates are of bounded variation and can be dominated as in Section 4.3, we can construct this model using our technology.

4.5 Proofs

In this section we will eventually construct a Markov process $(\omega_t)_{t \geq 0}$ meeting the conditions of Theorem 4.2.1. As mentioned in Section 4.2, we do this

by constructing the monotone political system with red and blue particles at each site, and then defining the original system as the difference between the two. The construction of the monotone political system is similar to the construction in Andjel [1].

Suppose we have a set $I \subset \mathbb{Z}^{\mathbb{Z}}$ and a nice collection of rate functions $\mathcal{F}_I$. This means that for every $i \in I$ we have a function f_i in the class $\mathcal{F}$, with corresponding functions f_i^r and f_i^b and constants c_i. We use these to give an informal description of the political process on $\Omega^+ \times \Omega^+$. The evolution of this process is slightly more complicated than that of the original process. The state of the political system at time t with initial configuration (x, y) is denoted by (x_t, y_t), where the number of red particles is given by x_t and the number of blue by y_t. For each configuration (x, y), position $j \in \mathbb{Z}$ and $i \in I$ we define the following addition rates:

$$\begin{aligned} R_{i,j}^{rb}(x,y) &= f_i^b\left(T^{-j}(x,y)\right) + c_i^-, \\ R_{i,j}^{r}(x,y) &= f_i^r\left(T^{-j}(x,y)\right) + c_i^+ - R_{i,j}^{rb}(x,y). \end{aligned}$$

The formal generator of the political process is then, for functions $g : \Omega^+ \times \Omega^+ \to \mathbb{R}$,

$$\begin{aligned} G^p g(x,y) = & \\ \sum_{i\in I, j\in\mathbb{Z}} & R_{i,j}^{rb}(x,y)\left(g(x + T^j i^+ + T^j i^-, y + T^j i^+ + T^j i^-) - g(x,y)\right) \\ & + R_{i,j}^{r}(x,y)\left(g(x + T^j i^+, y + T^j i^-) - g(x,y)\right). \end{aligned}$$

We can interpret this as follows. Particles are added according to $T^j(i^+, i^-)$ with rate $f^r(T^{-j}x, T^{-j}y) + c_i^+$, and according to $T^j(i^-, i^+)$ with rate $f^b(T^{-j}x, T^{-j}y) + c_i^-$. However, as many as possible of these births occur at the same times. The rate $R_{i,j}^r(x, y)$ therefore corresponds to the addition of just $T^j(i^+, i^-)$ and $R_{i,j}^{rb}(x, y)$ to the addition of $T^j(i^+ + i^-, i^+ + i^-)$. Notice that $R_{i,j}^r(x, y)$ is always positive. This follows from the fact that $f_i(x, y) = (c_i^+ + f_i^r(x, y)) - (f_i^b(x, y) + c_i^+) \geq 0$.

We will give a formal construction of the political process. Our strategy will be as follows. We first define a series of so called *political n-processes*, for $n \in \mathbb{N}$. Each of these processes will involve only a finite number of particles up to each time t and will be easy to define. Then we make a coupling of these processes, such that the limit of the (coupled) political n-processes is well defined. We define the political process as this limit.

Definition of the political n-processes
Let $n \in \mathbb{N}$ be fixed, let $I \subset \mathbb{Z}^{\mathbb{Z}}$ and $\mathcal{F}_I$ be a nice collection of rate functions, and write

$$r_{n,i,j}^r(x,y) := f_i^r(T^{-j}(x,y)) + c_i^+ \mathbb{1}_{\{j\in[-n,n]\}},$$

and

$$r^b_{n,i,j}(x,y) := f^b_i(T^{-j}(x,y)) + c^-_i \mathbb{1}_{\{j\in[-n,n]\}},$$

The political n-process with initial configuration $(x,y) \in \Omega^+ \times \Omega^+$ is defined as follows.

We start with some configuration of red and blue particles, but at time 0, all particles at sites in $\mathbb{Z} \setminus [-n,n]$ are removed. Then the political n-process evolves according to the following dynamics. Suppose the system is in state (u,v). For all $j \in \mathbb{Z}, i \in I$ we add $T^j(i^+ + i^-, i^+ + i^-)$ at rate

$$r^b_{n,i,j}(u,v),$$

and $T^j(i^+, i^-)$ at rate

$$r^r_{n,i,j}(u,v) - r^b_{n,i,j}(u,v).$$

We denote the state of the political n-process at time t and with initial configuration (x,y) by $(x_t,y_t)_n$ or $(x_{n,t}, y_{n,t})$. Observe that this description indeed gives rise to a process on $\Omega^+ \times \Omega^+$ since the expectation of the total number of particles in the political n-process is finite for any time $t > 0$.

Coupling of the political n-processes

We write $\mathbb{N}^+ = \{1,2,\ldots\}$. Let for $j \in \mathbb{Z}$, $i \in I$, $l \in \mathbb{N}^+$, $X_{i,j,l}(t)$ be independent Poisson processes with parameter 1, with an independent standard uniform random variable attached to each birth. We denote the state space of these Poisson processes and uniform random variables by Ω_p and write ρ for the associated measure. Typically, we denote an element of Ω_p by χ.

We use these processes to define a coupling $((\hat{x}_t, \hat{y}_t)_n)_{n\in\mathbb{N}}$ of the processes $(x_t,y_t)_{n,t\geq 0}$, $n \in \mathbb{N}$. In this coupling, when something is added in a political n-process, it is also added in all political m-processes with $m \geq n$. To be precise, in the political n-process, if the configuration at a certain moment is (u,v), for $j \in \mathbb{Z}$, $i \in I$ we add $T^j(i^+, i^-)$ if a Poisson arrival occurs in one of the processes

$$X_{i,j,1}(t), \ldots, X_{i,j,\lfloor r^r_{n,i,j}(u,v)\rfloor}(t),$$

or if a Poisson arrival occurs in the process

$$X_{i,j,\lceil r^r_{n,i,j}(u,v)\rceil}(t),$$

and the associated uniform random variable is less than

$$r^r_{n,i,j}(u,v) - \lfloor r^r_{n,i,j}(u,v)\rfloor.$$

Similarly, we add $T^j(i^-, i^+)$ if a Poisson arrival occurs in one of the processes

$$X_{i,j,1}(t), \ldots, X_{i,j,\lfloor r^b_{n,i,j}(u,v)\rfloor}(t),$$

or if a Poisson arrival occurs in the process

$$X_{i,j,\lceil r^b_{n,i,j}(u,v)\rceil}(t),$$

and the associated uniform random variable is less than

$$r^b_{n,i,j}(u,v) - \lfloor r^b_{n,i,j}(u,v)\rfloor.$$

Observe that we use the same set of Poisson processes and associated uniform random variables for all political n-processes and to decide whether we add $T^j(i^+, i^-)$ or $T^j(i^-, i^+)$. This ensures that $T^j(i^+ + i^-, i^+ + i^-)$ and $T^j(i^+, i^-)$ are added with the required rates. In the coupling described above, we denote the state of the political n-process at time t and with initial configuration (x, y) by $(\hat{x}_t, \hat{y}_t)_n$, or $(\hat{x}_{n,t}, \hat{y}_{n,t})$. Sometimes, if we want to stress the dependence on the realisation of the Poisson and uniform processes, we write $(\hat{x}_t, \hat{y}_t)_n(\chi)$.

Observe that by the monotonicity of the rates (both in the configuration and in the parameter n) we have that if $(x, y) \leq (u, v)$ then $(\hat{x}_t, \hat{y}_t)_n \leq (\hat{u}_t, \hat{v}_t)_n$, and that for $n \geq 1$, $(\hat{x}_t, \hat{y}_t)_n \leq (\hat{x}_t, \hat{y}_t)_{n+1}$.

Definition of the political process

We define the political process with initial configurations in $\Omega^+ \times \Omega^+$ and configurations at time $t > 0$ in $(\mathbb{N} \cup \{\infty\})^{\mathbb{Z}} \times (\mathbb{N} \cup \{\infty\})^{\mathbb{Z}}$. Let $(x, y) \in \Omega^+ \times \Omega^+$. We use the coupling described above to define $(x_t, y_t) \in (\mathbb{N} \cup \{\infty\})^{\mathbb{Z}} \times (\mathbb{N} \cup \{\infty\})^{\mathbb{Z}}$ by

$$(x_t, y_t) := \lim_{n\to\infty} (\hat{x}_t, \hat{y}_t)_n$$

This is well defined, because of the monotonicity of the n-processes in the parameter n. We define $(x_t, y_t)_{t\geq 0}$ to be the political process.

We next prove that if this process starts in a configuration such that the expected number of particles per site is bounded, then this is also true at any later time.

Lemma 4.5.1 *Let $\tilde{\nu}$ be a stationary ergodic measure on $\mathbb{N}^{\mathbb{Z}} \times \mathbb{N}^{\mathbb{Z}}$, with*

$$\int_{\mathbb{N}^{\mathbb{Z}}\times\mathbb{N}^{\mathbb{Z}}} (x(0) + y(0))\, d\tilde{\nu}(x, y) < \infty.$$

Let $I \subset \mathbb{Z}^{\mathbb{Z}}$, let $\mathcal{F}_I$ be a nice collection of rate functions and let $\mu = \tilde{\nu} \times \rho$. Then for all $q \in \mathbb{Z}$,

$$\int_{(\Omega^+\times\Omega^+)\times\Omega_p} (x_t(q) + y_t(q))\, d\mu < \infty,$$

and

$$\mu\left(((x, y), \chi) : (x_t, y_t) \in \Omega^+ \times \Omega^+, \forall t \geq 0\right) = 1.$$

Proof of Lemma 4.5.1 We first sketch the idea of the proof. We will make a comparison with another process, which we call the *growth process* (to be defined), with initial configurations in $\Omega^+ \times \Omega^+$ and configurations at time t in $((\mathbb{N} \cup \{\infty\}) \times (\mathbb{N} \cup \{\infty\}))^{\mathbb{Z}}$. We make a coupling of this growth process and the political process, with the property that, at every site, the number of red and blue particles in the political process is not larger than the number of red and blue particles at the same site in the growth process. Then we show that, under a suitable initial measure, the growth process will always stay in $\Omega^+ \times \Omega^+$, almost surely, and the expected number of particles per site will stay finite. We do this by reinterpreting the growth process as a spatial branching process. The growth process dominates the political process, which leads to the desired result for the political process.

We begin by giving an informal description of the growth process. In this process, for all $j \in \mathbb{Z}$, $i \in I$ and $\xi = (\xi^1, \xi^2) \in \Omega^+ \times \Omega^+$, $T^j(||i||, ||i||)$ is added at rate

$$k_i \cdot (T^{-j}(\xi^1 + \xi^2)) + |c_i| = \sum_{l \in \mathbb{Z}} k_i(l)(\xi_1(l-j) + \xi_2(l-j)) + |c_i|.$$

The formal definition of the growth process proceeds via a limit of growth n-processes and is completely analogous to the formal definition of the political process. We leave the details to the reader.

If we let the transitions in the growth process be governed by the same $X_{i,j,l}(t)$ and associated uniform random variables as in the political process, we obtain a coupling of the growth process and the political process. In this coupling, we denote the state of the growth process at time t and with initial configuration ξ by ξ_t. We find that in this coupling, if the growth process starts with the same number of particles per site as the political process, for each site the number of particles in the political process is at most the number of particles at the corresponding site in the growth process. In formulas, if $(x, y) \in \Omega^+ \times \Omega^+$ and if $\xi = (x, y)$, then $(x_t, y_t) \leq \xi_t$.

To prove the lemma, it suffices to show that for all $q \in \mathbb{Z}$ and for all $t > 0$,

$$\int_{(\mathbb{N}^{\mathbb{Z}} \times \mathbb{N}^{\mathbb{Z}}) \times \Omega_p} (x_t(q) + y_t(q))\, d\mu < \infty. \tag{4.1}$$

To prove (4.1) we use the coupling with the growth process mentioned above, and we can show that

$$\mathbb{E}_\mu(\xi_t^1(q) + \xi_t^2(q)) := \int_{\mathbb{N}^{\mathbb{Z}} \times \Omega_p} \xi_t^1(q) + \xi_t^2(q)\, d\mu < \infty, \tag{4.2}$$

we are done. To prove (4.2), we explain how the growth process can be interpreted as a spatial branching process with immigration. For all $j \in \mathbb{Z}$,

and $i \in I$, $T^j(||i||, ||i||)$ is added at rate $|c_i|$. This is the immigration part of the process. Further each particle at site j gives birth to children distributed according to $T^j(||i||, ||i||)$ at rate $k_i(j-l)$, for each $i \in I$, $l, j \in \mathbb{Z}$. Notice that the expected birth rate per particle is $\kappa = \sum_{i \in I} 2|k_i||i| < \infty$, and the expected immigration rate per site is $\gamma < \infty$.

We can now find an upper bound for the expected number of particles at a site, in the growth process, if the initial measure is $\tilde{\nu}$. Write

$$\lambda := \int_{\mathbb{N}^{\mathbb{Z}}} (\xi^1(0) + \xi^2(0))\, d\tilde{\nu}(\xi) < \infty.$$

Define the earliest ancestor of a particle to be either the particle at time 0 that has the particle as descendent, or the earliest immigrant ancestor of the particle. Let $v_t(j,l)$ be the number of particles at position j at time t which had an ancestor at position l. Noting that the expected number of immigrants per site up to time t is γt and using the (spatially) stationarity of the branching process, we see that the expected number of particles in the growth process at time t and position q, $\xi_t^1(q) + \xi_t^2(q)$, satisfies:

$$\begin{aligned} \mathbb{E}_\mu(\xi_t^1(q) + \xi_t^2(q)) &= \sum_{l=-\infty}^{\infty} \mathbb{E}_\mu[v_t(q,l)] \\ &= \sum_{l=-\infty}^{\infty} \mathbb{E}_\mu[v_t(l,q)] \\ &\leq (\lambda + \gamma t) e^{\kappa t} < \infty. \end{aligned}$$

□

We are nearly ready to prove Theorem 4.2.1, but to prove the relationship between $S(t)$ and G we will need the corresponding equation for the political process, and for this we will need the same connection in the growth process. These relations are given in the following two lemmas, the somewhat technical proofs of which can be found after the proof of Theorem 4.2.1.

Lemma 4.5.2 *Let $h^*(\xi^1, \xi^2) : \Omega^+ \times \Omega^+ \to \mathbb{R}_{\geq 0}$ be a monotone increasing function such that $h^*(\xi^1 + \mathbf{1}_j, \xi^2) - h^*(\xi^1, \xi^2), h^*(\xi^1, \xi^2 + \mathbf{1}_j) - h^*(\xi^1, \xi^2) \leq z(j)$, for some z with $|z| \leq \infty$. Let G^g and S^g be the formal generator and semigroup for the growth process and let $\tilde{\nu}$ be a stationary, ergodic measure on $\Omega^+ \times \Omega^+$ such that*

$$\int_{\Omega^+ \times \Omega^+} (\xi^1(0) + \xi^2(0))\, d\tilde{\nu}(x,y) < \infty.$$

Let $\mu = \tilde{\nu} \times \rho$. Then,

$$\mu\left(\lim_{m\to 0} \frac{S^g(m)h^*(\xi_t^1,\xi_t^2) - h^*(\xi_t^1,\xi_t^2)}{m} = G^g h^*(\xi_t^1,\xi_t^2) \text{ for all } t\right) = 1.$$

Lemma 4.5.3 *Let $h^*(x,y) : \Omega^+ \times \Omega^+ \to \mathbb{R}_{\geq 0}$ be a monotone increasing function such that $h^*(x+\mathbf{1}_j,y) - h^*(x,y), h^*(x,y+\mathbf{1}_j) - h^*(x,y) \leq z(j)$, for some z with $|z| \leq \infty$. Let G^p and S^p be the formal generator and semigroup for the political process and let $\tilde{\nu}$ be a stationary, ergodic measure on $\Omega^+ \times \Omega^+$ such that*

$$\int_{\Omega^+\times\Omega^+} (x(0)+y(0))\, d\tilde{\nu}(x,y) < \infty.$$

Let $\mu = \tilde{\nu} \times \rho$. Then,

$$\mu\left(\lim_{m\to 0} \frac{S^p(m)h^*(x_t,y_t) - h^*(x_t,y_t)}{m} = G^p h^*(x_t,y_t) \text{ for all } t\right) = 1.$$

Proof of Theorem 4.2.1 Let ν, I and $\mathcal{F}_I$ and h be as in the theorem. Let f_i^r and f_i^b be the associated monotone rate functions, for all $i \in I$. For $\omega \in \Omega$, define $(x,y) := (\omega^+, \omega^-)$ and consider the political process with initial configuration (x,y). Let $\tilde{\nu}$ be the law of (x,y), if ω is chosen according to ν. Observe that $\tilde{\nu}$ is ergodic, as it is a factor of an ergodic measure, and that

$$\int_{\Omega^+\times\Omega^+} (x(0)+y(0))\, d\tilde{\nu}(x,y) < \infty.$$

By Lemma 4.5.1 we find

$$(\nu\times\rho)\left((\omega,\chi) \in \Omega\times\Omega_p : x_t, y_t \in \mathbb{N}^{\mathbb{Z}}, \forall t\right) = 1.$$

We define $\omega_t := x_t - y_t$ if this exists in Ω and $\omega_t = \flat$ otherwise, and observe that

$$(\nu\times\rho)\left((\omega,\chi) : \omega_t \neq \flat, \forall t \geq 0\right) = 1.$$

Note that ν^* is the marginal measure of $\nu\times\rho$ for the original process $(\omega_t)_{t\geq 0}$.

Given Lemma 4.5.3, the proof of the theorem follows fairly quickly. We can write

$$h(x-y) = h^r(x,y) - h^b(x,y)$$

where h^r and h^b are monotone increasing, positive functions such that for every $(x,y) \in \Omega^+ \times \Omega^+$,

$$h^*(x+\mathbf{1}_j,y) - h^*(x,y), h^*(x,y+\mathbf{1}_j) - h^*(x,y) \leq z(j), \text{ for } * = r, b.$$

Note that, for $\omega \in \Omega$,

$$\begin{aligned}\frac{S(m)h(\omega_t) - h(\omega_t)}{m} &= \frac{S^p(m)h(x_t - y_t) - h(x_t - y_t)}{m} \\ &= \frac{S^p(m)h^r(x_t, y_t) - h^r(x_t, y_t)}{m} \\ &\quad - \frac{S^p(m)h^b(x_t, y_t) - h^b(x_t, y_t)}{m}.\end{aligned}$$

Lemma 4.5.3 tells us that, for all t, the last part of this equation converges to

$$G^p h^r(x_t, y_t) - G^p h^b(x_t, y_t) = G^p h(x_t - y_t) = Gh(\omega_t)$$

as m tends to zero, with μ-probability 1, by the definition of the original process in terms of the political process.

Notice that all the rates are "correct". This follows from fact that if both red and blue particles appear at the same time then this is not seen in the difference, and the careful choice of rates. For example, the rate with which we add only (i^+, i^-) is,

$$f_i^r(x, y) + c_i^+ - (f_i^b(x, y) + c_i^-) = f(x - y)$$

□

Proof of Lemma 4.5.2 Notice that,

$$\begin{aligned}0 &\leq \frac{S^g(m)h(\xi_t^1, \xi_t^2) - h(\xi_t^1, \xi_t^2)}{m} - G^g h(\xi_t^1, \xi_t^2) \\ &\leq \frac{S^g(m)(z \cdot (\xi_t^1 + \xi_t^2)) - z \cdot (\xi_t^1 + \xi_t^2)}{m} - G^g(z \cdot (\xi_t^1 + \xi_t^2)), \quad (4.3)\end{aligned}$$

by the domination of any increase in h at site j by $z(j)$. This last term will be dominated by the corresponding expression for a simpler process on $\mathbb{R}_{\geq 0}$, the *simple process*, which we define informally as follows.

Let $\theta(j) = k(j) + z(j)$, and notice that $\phi = \sum_{j \in \mathbb{Z}} \theta(j) < \infty$. Let $\alpha_{(i,j)} = \theta \cdot T^j(2||i||)$, and notice that $\alpha = \sum_{i \in I, j \in \mathbb{Z}} \alpha_{(i,j)} < \infty$. Notice also that $\sigma = \sum_{j \in \mathbb{Z}, i \in I} \alpha_{(i,j)} c_i < \infty$. Let a be a positive real number. At rate $a + c_i$, a changes to $a + \alpha_{(i,j)}$, for all $j \in \mathbb{Z}, i \in I$. We can define this process using its monotonicity as for the political process. Let G^s be the generator of the simple process and S^s the semigroup. Then we claim that

$$\lim_{m \downarrow 0} \frac{S^s(m)a - a}{m} = G^s a.$$

To prove this claim we first define an n-process, a_t^n. This allows additions of $\alpha_{i,j}$ only if $j \in [-n, n]$ and i is in the first n elements of some fixed listing

of $I = \{i_1, i_2, \ldots\}$. We will define $a_t = \lim_{n\to\infty} a_t^n$. Let the generator and semigroup of the n-process be given by G_n^s and $S_n^s(m)$. Notice that

$$\lim_{m\downarrow 0} \frac{S_n^s(m)a - a}{m} = G_n^s(a),$$

and that

$$G_n^s(a) = \sum_{j\in[-n,n],l=1,\ldots,n} \alpha_{(i,j)}(a + c_{i_l}) \leq \alpha a + \sigma.$$

Therefore

$$S_n^s(m)a_t = \sum_{p=0}^{\infty} \frac{m^p (G_n^s)^p a}{p!}.$$

Observe that by monotone convergence

$$\begin{aligned} S^s(m)a &= \lim_{n\to\infty} \sum_{p=0}^{\infty} \frac{m^p (G_n^s)^p a}{p!} \\ &= \sum_{p\in\mathbb{Z}} \frac{m^p (G^s)^p a}{p!} \end{aligned}$$

by dominated convergence. The claim follows. Note that a_t is finite for all t, almost surely.

The simple process can be coupled to $z \cdot (\xi_t^1 + \xi_t^2)$. We begin with a given by $\theta \cdot (\xi_1 + \xi_2)$, where (ξ_1, ξ_2) is distributed according to $\tilde{\nu}$. Note that $a < \infty$ μ-almost surely. We couple additions of $T^j(||i||, ||i||)$ in the growth process with additions of $\alpha_{i,j}$ in the simple process. This means that we always add more to a than $z \cdot (\xi_t^1 + \xi_t^2)$, and at a faster rate, because $\theta(j) \geq z(j)$ and $\theta(j) \geq k(j)$ respectively. Therefore, for all t,

$$\mu(\lim_{m\downarrow 0} \frac{S^s(m)a_t - a_t}{m} = G^s a_t \text{ for all } t) = 1$$

as a_t is μ-almost surely finite and well defined.

By the coupling, for all t,

$$\frac{S^g(m)h(\xi_t^1, \xi_t^2) - h(\xi_t^1, \xi_t^2)}{m} - G^g h(\xi_t^1, \xi_t^2) \leq \frac{S^s(m)a_t - a_t}{m} - G^s a$$

μ-almost surely. This combined with Equation 4.3 proves the lemma. □

Proof of Lemma 4.5.3 Note that by our coupling, for all t,

$$\begin{aligned} 0 &\leq \frac{S^p(m)h^*(x_t, y_t) - h^*(x_t, y_t)}{m} - G^p h^*(x_t, y_t) \\ &\leq \frac{S^g(m)h^*(\xi_t^1, \xi_t^2) - h^*(\xi_t^1, \xi_t^2)}{m} - G^g h^*(\xi_t^1, \xi_t^2) \end{aligned}$$

μ-almost surely. By Lemma 4.5.2 we know that the last term converges to 0 as m tends to zero, with μ-probability one. Lemma 4.5.3 follows immediately. □

4.6 Remarks

In this section we mention a couple of possible extensions to this theory and outline an alternative approach, suggested by Bálint Tóth.

- In this article we have been able to define a process by defining it as the difference of two monotone coupled processes. We could do this as both of these processes were shown to have finitely many particles per site at any given time, almost surely. We were able to prove this by comparison with a branching process. We cannot make this comparison if the rate functions grow faster than linearly but we conjecture that if both the processes grow at most polynomially then the same construction should work. In this case we would have to have to restrictions upon higher moments of $w(0)$ in order that the growth rate be finite.

- It would also be possible to use this construction technique if only one of the processes was constrained to be finite. We would then have the possibility of having ∞ or $-\infty$ particles at some sites, but not both.

- While we have here defined our process as the difference of two processes, it might be possible to define certain other processes as some other function of two (or more) processes.

An Alternative Approach
The alternative approach to the construction problem involves an alternative political process. We again set $(x_0, y_0) = (\omega^+, \omega^-)$ and but now add $T^j(i^+, i^-)$ to (x, y) with rate $f_{i^+,i^-}(T^{-j}(x, y)) = f_i(T^{-j}(x - y))$. As in the earlier approach, if we can define this process we can define our original process by $w_t = x_t - y_t$. However the rates in the political process are no longer necessarily monotone.

We can still dominate this alternative political process by our growth process, and this means that we can "thin" the n-growth processes to give alternative n-political processes. Up to any time t, and for any block $[-m, m]$, there almost surely exists a finite block (of random size), which contains all ancestors of the particles in $[-m, m]$ in the growth process. This means that we can define the alternative political process as the limit of the alternative n-processes, and thus retrieve the original process. The relationship between semigroup and generator still holds.

We thank Ronald Meester and Marton Balàzs.

Bibliography

[1] E. D. Andjel (1982), Invariant measures for the zero range process, *The Annals of Probability*, 10(3), 525-547.

[2] M. Balàzs (2001), Microscopic shape of shocks in a domain growth model, *Journal of Statistical Physics*, 105(3/4), 511-524.

[3] T.M. Liggett (1980), Long range exclusion processes, *The Annals of Probability*, 8(5), 861-889.

[4] T.M. Liggett (1985), *Interacting Particle Systems*, Springer.

[5] T.M. Liggett (1999), *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*, Springer.

[6] C. Maes, F. Redig, E. Saada and A. Van Moffaert (2000), On the thermodynamic limit for a one-dimensional sandpile process, *Markov Processes and Related Fields*, 6(1), 1-21.

[7] R. Meester and C.M. Quant (2001), On a long range particle system with unbounded flip rates, *preprint*.

Samenvatting

Dit proefschrift bevat drie artikelen over drie verschillende modellen. Alle artikelen gaan over kansrekening in oneindig grote ruimten. Oneindig grote systemen lijken vaak ingewikkeld. Toch is het mogelijk intuïtie voor dit soort systemen op te bouwen zonder alle technische details te begrijpen. Hier volgt een globale uitleg van de modellen met behulp van interactieve raadsels en een samenvatting van de resultaten.

$\infty \quad \infty \quad \infty \quad \infty$

Een realisatie van het eerste model is te zien in Figuur 1.1. Voor elke cirkel die je ziet, heb ik een eerlijke munt opgegooid. Staat in de cirkel $+1$, dan kwam kop boven, bij -1 kwam munt boven. In de figuur kunnen we over het lichtgrijze pad van S naar F lopen. Tijdens onze wandeling tellen we de getallen op die we tegenkomen. Alle uitkomsten van de optellingen (de *partiële sommen*) langs het pad liggen tussen -5 en $+2$, ofwel in het *interval* $[-5, 2]$. Als we langs het donkergrijze pad lopen, zien we dat de partiële sommen in het kortere interval $[-3, 3]$ liggen. De twee grijze paden worden *zichzelf ontwijkende paden* genoemd, omdat je iedere cirkel hoogstens één keer tegenkomt als je over het pad loopt.

Raadsel Bestaat er een zichzelf ontwijkend pad van S naar F in Figuur 1.1 met partiële sommen in het interval $[-1, 1]$? Is er een pad met partiële sommen in een kleiner interval?

Vervolgens bekijken we Figuur 1.2. Om deze figuur te maken, heb ik een oneerlijke munt opgegooid waarbij de kans dat kop bovenkomt (en dus $+1$) $6/7$ is.

Raadsel Voor welk zichzelf ontwijkend pad van S naar F in Figuur 1.2 liggen alle partiële sommen in een zo klein mogelijk interval?

In Hoofdstuk 2 bestuderen we hetzelfde soort problemen, maar dan op een oneindig groot rooster van cirkels. Daarbij vragen we ons af of er een oneindig lang zichzelf ontwijkend pad bestaat met sommen in een bepaald interval. Als de kans op kop dichtbij $1/2$ ligt, blijkt dat er inderdaad een

oneindig lang pad kan zijn met sommen in een zeker interval, namelijk in het interval $[-21, 21]$. Als de kans op kop dicht bij 1 ligt, bestaat er echter geen interval waarin de sommen zullen blijven. Dit soort fenomenen noemt men een *fase-overgang*. Deze resultaten gelden ook voor zogenaamde *gerichte paden* waarbij alle stappen omhoog of naar rechts genomen worden.

We onderzoeken ook of er paden bestaan waarin de sommen oneindig vaak nul zijn; hierbij treedt ook een fase-overgang op. Verder tonen we een verbazingwekkend verband aan tussen het probleem om een pad naar oneindig te vinden (het pad hoeft niet zelf-ontwijkend te zijn, maar mag ieder roosterpunt slechts eindig vaak bezoeken) met partiële sommen in een bepaald interval en het klassieke Boolese model met vierkanten om de punten van een zogenaamd Poisson proces in het vlak.

$\infty \ \infty \ \infty \ \infty$

Telecommunicatie-bedrijven kiezen waar ze hun zendmasten plaatsen aan de hand van de locaties van hun klanten. We modelleren dit proces in Hoofdstuk 3 en beschouwen de gevolgen van locale regels. In Figuur 1.4 zie je een aantal klanten (punten) en ontvangstgebieden van zendmasten (schijven). Een *overdekking* met zendmasten voldoet aan twee eisen. Ten eerste moet elke klant zich binnen het ontvangstgebied van een zendmast bevinden. Ten tweede moet elke zendmast een klant binnen zijn ontvangstgebied hebben. Een voorbeeld van een overdekking is te zien in Figuur 1.4.

Raadsel Hoeveel zendmasten heb je minimaal nodig om de klanten in Figuur 1.4 en Figuur 1.7 te overdekken?

Zendmasten kunnen met elkaar communiceren als hun ontvangstgebieden overlappen. Hierdoor kunnen klanten communiceren over een langere afstand, zoals in Figuur 1.5 te zien is.

Raadsel Kun je de klanten in Figuur 1.6 overdekken, zodat Elm Cottage met Feverfew House kan communiceren? Hoeveel zendmasten heb je minimaal nodig? Kun je een overdekking maken zodat Elm Cottage *niet* met Feverfew House kan communiceren? Kun je dit doen met slechts 4 zendmasten? Is het mogelijk de klanten in Figuur 1.7 te overdekken zodat Gorse Cottage berichten aan Hazel House kan sturen?

In Hoofdstuk 3 nemen we aan dat de klanten een willekeurige locatie op een oneindig grote aarde hebben, of wiskundig gezegd, dat de klanten volgens een Poisson proces in het vlak geplaatst worden. We bekijken nog steeds hetzelfde soort problemen. Als er een oneindig grote groep klanten is die allemaal met elkaar kunnen communiceren, noemen we dit een *oneindige tros*. Als de dichtheid mensen heel klein is, laten we zien dat er geen oneindige tros

kan zijn. Maar we laten ook zien dat als de dichtheid klanten heel groot is, er niet automatisch een oneindige tros hoeft te bestaan. Dit geldt ook als de manier waarop we de zendmasten plaatsen *onafhankelijk van verschuivingen* is. Dat wil zeggen dat als alle klanten over een zekere afstand in een bepaalde richting verschuiven, de posities van de zendmasten in de nieuwe overdekking precies de verschoven posities van de oorspronkelijke overdekking zijn.

Als alle zendmasten echter op een roosterpunt geplaatst moeten worden en de dichtheid van klanten groot genoeg is, bestaat er wel een oneindige tros. Dit is ook zo in een *n-vierkant overdekking*, waarbij het vlak opgedeeld is in vierkanten van zijde n en we in elk van deze vierkanten de minimaal mogelijke hoeveelheid zendmasten plaatsen. De laatste twee soorten overdekkingen zijn voorbeelden van zogenaamde *platte overdekkingen* waarbij er een maximale hoeveelheid zendmasten per vierkante meter is. Voor alle platte overdekkingen wordt er een oneindige tros gevormt als de dichtheid klanten groot genoeg is.

We bekijken ook het geval waarin de zendmasten met elkaar kunnen communiceren als hun onderlinge afstand kleiner of groter is dan twee maal de straal van de ontvangstgebied. Als deze afstand groter is, moet er een oneindige tros bestaan als de dichtheid klanten maar groot genoeg is.

Met n-vierkant overdekkingen kunnen we willekeurig dicht bij de minimaal benodigde dichtheid zendmasten komen. Er bestaat zelfs een overdekkings-algoritme met deze minimale dichtheid.

$\infty \quad \infty \quad \infty \quad \infty$

Het laatste artikel gaat over het bestaan en de definitie van oneindige deeltjes-systemen. Waarom het nodig is om over het bestaan van zulke systemen na te denken, kan misschien het beste uitgelegd worden aan de hand van een voorbeeld van een oneindig systeem dat niet bestaat.

Een rij van k eerlijke munten ligt op tafel. Elke munt heeft een zwarte en een witte kant. Om te beginnen gooien we alle munten op. Een mogelijke uitkomst is te zien in Figuur 1.8. Munt 1 bevindt zich in een blokje van 3 zwarte munten. Deze drie zwarte munten veranderen van kleur na een stochastische tijd van gemiddeld $\left(\frac{1}{2}\right)^3 = \frac{1}{8}$ seconde. In het algemeen verandert een blok van n munten van dezelfde kleur, van kleur na gemiddeld $\left(\frac{1}{2}\right)^n$ seconde, onafhankelijk van wat er gebeurt met andere blokken. Een blok van dezelfde kleur groeit als een van de buur-blokken van kleur verandert, of als het zelf van kleur verandert. Het is mogelijk te laten zien dat, als we aan het begin een blok van grootte n hebben, het gemiddelde niet langer dan 2×2^{-n} seconde duurt totdat het hele systeem dezelfde kleur heeft. Met andere woorden: alle munten worden dezelfde kleur na een tijd die korter

wordt naar mate het grootste blok groter wordt. Voor eindige k is dit geen probleem. Als er echter oneindig veel munten zouden zijn, zouden we op tijdstip 0 blokjes van elke lengte op tafel zien. Dit betekent dat op het ogenblik dat we beginnen alles direct dezelfde kleur zou moeten worden. Dat is een heel ander gedrag dan hierboven beschreven werd.

In het laatste hoofdstuk geven we een aantal condities waaronder zulke rare verschijnselen gelukkig niet optreden. Wiskundig gezegd, we construeren deeltjes-systemen op een rooster, $\mathbb{Z}^d$, waarbij er geen beperking is op het aantal deeltjes per roosterpunt. Deze systemen beginnen in configuraties gekozen volgens een stationaire, ergodische maat, waarbij het verwachte aantal deeltjes per roosterpunt eindig is. Deeltjes worden toegevoegd en weggenomen met snelheden die functies van de configuraties zijn. Deze functies zijn van begrensde variatie en zijn zo dat de snelheid waarmee deeltjes toegevoegd of weggenomen worden eindig is. Negatieve aantallen deeltjes zijn toegestaan.

Acknowledgements

Many thanks to:

Rob van den Berg, Michel Dekking, Richard Gill, Harry Kesten, and Bálint Tóth, for carefully reading my thesis,

Ronald Meester, for being the perfect supervisor - by inspiring me, encouraging me and meticulously placing full stops after my formulae,

Corrie Quant for infinite patience with my Dutch, mathematics and piles of paper,

everyone at the mathematics department in Utrecht for help, support and conversations, and in particular Martijn Pistorius for help with the samenvatting,

Massimo Franceschetti for sharing the problem that was to become Chapter 3 of this thesis,

my hosts at T.I.F.R. in Mumbai, I.S.I. in Delhi and at the Technical University in Budapest for very enjoyable visits and for making me feel thoroughly welcome,

my friends in the Netherlands and abroad, especially Karin, Derek, Federica, Luciano, Marketa and my parents,

everybody at the Tussenbus, for being real people,

Dave, for being a very, very nice chap indeed,

and to everyone else who should, could or would secretly like to be thanked.

Curriculum Vitae

Lorna Booth was born on 23rd December 1976 in Stockport, England and grew up in the village of Quorn in Leicestershire. She went on to study mathematics at King's College in Cambridge University, and then came to Utrecht University in 1998 to write a Ph.D. about probability. Her hobbies include serving coffee on a bus and learning things.