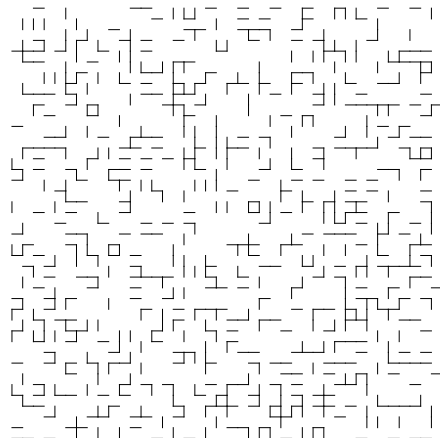
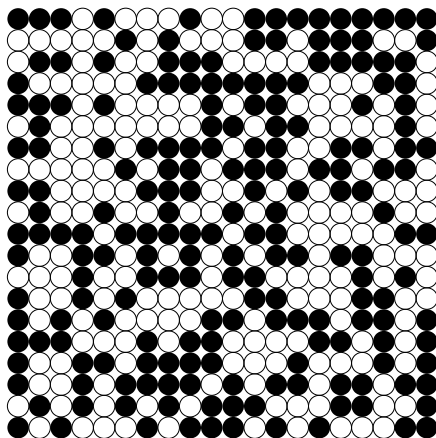
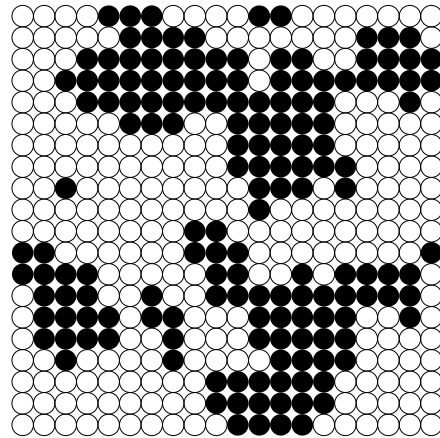
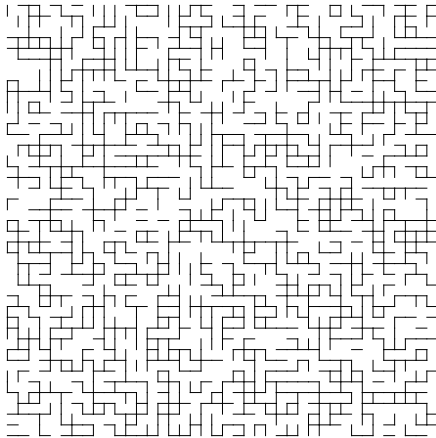


Probabilistic Methods in Statistical Physics

Thierry Lévy *

Tsinghua University, October 2008



*CNRS and École Normale Supérieure – DMA – 45, rue d'Ulm – F-75005 Paris, levy@dma.ens.fr

Contents

I	Percolation	7
1	Binomial Galton-Watson trees	7
1.1	Galton-Watson processes	7
1.2	A tree with branches removed at random	8
2	The percolation process	9
2.1	Introduction	9
2.2	Graphs and edge configurations	10
2.3	The measurable space of percolation	11
2.4	The percolation measure	12
2.5	Order and increasing events	13
2.6	Approximation results	14
2.7	Coupling the percolation processes	15
2.8	The phase transition phenomenon	16
3	Correlation inequalities and stochastic ordering	19
3.1	The FKG inequality	19
3.2	The BK inequality	21
3.3	Stochastic ordering	21
4	The cluster of the origin and the number of infinite clusters	24
4.1	Existence of an infinite cluster	24
4.2	The number of infinite clusters in the supercritical phase	25
4.3	The cluster of the origin in the subcritical phase	27
5	The critical probability in two dimensions	30
II	The Ising model	34
6	The probability measure	34
7	First examples	35
7.1	The segment	36
7.2	The complete graph	36
8	The phase transition	37
8.1	Magnetization	37
8.2	The one-dimensional case	38

9	The random cluster model	40
9.1	The random cluster measure	40
9.2	Inequalities between random cluster model and percolation	42
9.3	Coupling the random cluster and Ising models	43
10	Infinite volume limits	45
10.1	Boundary conditions	45
10.2	The random cluster model on \mathbb{L}^d	46

Introduction

Statistical physics is a branch of physics which aims at deducing macroscopic properties of large systems from the knowledge of the behaviour of its microscopic components. The words *macroscopic* and *microscopic* should not be taken too literally: stars can for instance be considered as the microscopic constituents of galaxies. Also, the same system can be micro- or macroscopic, depending on the problem under study: a droplet of water is a macroscopic system consisting of a huge number of (microscopic) molecules, and droplets of water are microscopic constituents of a (macroscopic) cloud. Finally, the word *deduce* may be misleading. Macroscopic theories have often been discovered before, or at least independently of microscopic ones. In these cases, one of the goals of statistical physics is to provide a bridge between two different approaches to the same phenomenon.

It may be surprising that it takes a theory to deduce macroscopic properties from microscopic ones. Once one has chosen a physical theory of fundamental interactions, the size of the systems under consideration should not be an issue. In the case of galaxies, one could choose to work with Newtonian mechanics, or with general relativity, but in both cases, the equations should be valid for 10^{11} stars (apparently the order of magnitude of the number of stars in the Milky Way) just as well as for two or three. In a droplet of water, the 10^{20} molecules could be modeled by electric dipoles to which Coulomb's law applies just as it does for two electrons. But in fact, anyone trying to actually compute anything on a system of 10^{20} molecules using Newtonian mechanics and classical electrostatics would be running into serious trouble. One should at least know the state of the system, that is, the positions, speeds, and orientations of the molecules. Just to be sure that this is far beyond what is technically possible, let us make a short computation. For each molecule, one should store $3 + 3 + 2 = 8$ real numbers. In a computer, this takes about 30 bytes. Now, assuming that one stores one gigabyte on a one gram chip, storing 10^{21} bytes would take a one million ton memory.

This illustrates the fact that even the best physical theory, which would perfectly describe the fundamental interactions between the elementary bricks of the universe, would not, without a serious amount of work, at the same time be an efficient theory of the everyday world, in which every system is incredibly complex.

One way to circumvent the complexity of real-world systems is to replace a perfect description, which is practically inaccessible, by a statistical description. For example, in a droplet of water, instead of trying to record the speed of each molecule, one can try to describe the statistical distribution of the various speeds: in other words to determine, for each possible speed, which proportion of the molecules have a speed which is close to this one. One could also do this for the positions, and for the orientations of the molecules. At a higher level of sophistication, one could ask for the joint distribution of the speed and orientation, that is, given a speed and an orientation, to ask for the proportion of molecules which have at the same time a speed and an orientation close to the specified ones.

It turns out that a lot can be said about macroscopic systems by treating their microscopic constituents as if they were behaving randomly, according to statistical rules

on which one makes appropriate assumptions. One of the first successes of this approach has been the kinetic theory of gases. Let us give another classical example. Imagine that we want to study the diffusion of a drop of coffee released, with the help of some clever device, in the middle of a bucket of clear water. The “molecules of coffee” are being shocked constantly and in all directions by the surrounding molecules of water. It seems reasonable to model their movement by a random walk or a Brownian motion. Thus, the distribution of the molecules of coffee at a certain time t should be close to that of a Brownian motion at time ct for some appropriate constant c , that is, a Gaussian distribution with variance ct . If $c = 1$, then the concentration of coffee at distance r from the original drop should thus behave like $\frac{1}{\sqrt{t}}e^{-\frac{r^2}{2t}}$. A simple probabilistic argument gives us a good first approximation of the behaviour of a complicated macroscopic system.

This explains the name of statistical physics, and why there should be probabilistic methods involved. However, the main subject of the present lectures is another aspect of statistical physics, namely phase transitions. Let us think of a small number of molecules of water in an empty box. Their motion is governed by rules which, mathematically, take the form of simple differential equations. One should expect that the various observable quantities attached to this system depend on each other in a very smooth way. This is certainly true for our small system. However, in a macroscopic system, this is not the case anymore. The fact that water freezes when one cools it down is an example: we could say that the viscosity of an assembly of molecules of water does not depend smoothly on the average kinetic energy of these molecules. And indeed the freezing of water is by no means a smooth transition: there is no intermediate state between water and ice. We can easily observe a mixture of both, but not something like sticky water or soft ice.

There is no real contradiction: we may still agree that the viscosity of water depends smoothly on temperature, because it must, on mathematical grounds, but it varies so rapidly around a certain critical value of the temperature that it is discontinuous for all practical purposes. If it was possible to consider an infinite system of molecules of water, then we would certainly observe a genuine discontinuity. Of course, no one will ever observe such an infinite system, but still, the idea is meaningful, as we shall see.

The mathematical challenge is thus to find simple probabilistic models which depend on a parameter and whose behaviour, around some critical value of this parameter, changes suddenly. More precisely, to find observable quantities of these models which depend in a non-smooth way on the parameter. By *non-smooth*, we mean: discontinuous, or not differentiable, or not twice differentiable, or not analytic... There is a classification of phase transitions according to the lack of regularity that is observed at the transition.

It is not so easy to design a probabilistic model with a phase transition. We will discuss two of them, percolation and the Ising model. Percolation was introduced in around 1960 to model the diffusion of water in porous rocks, and the Ising model was proposed by Lenz in 1924 in the context of ferromagnetism.

Here is a summary of the lectures. We start by a brief review of Galton-Watson processes and interpret the classical dichotomy on the mean of the offspring distribution as a phase transition. We observe also that Galton-Watson processes with binomial offspring distribution can be seen as percolation on a tree. We do not use any results of

this introductory section later. Then, we define and study percolation, on an arbitrary graph. Our goal in this first half of the lectures is to prove that there is a phase transition on the cubic lattice in any dimension larger than 2, and to prove at least a part of the result of Kesten which asserts that the critical probability in two dimensions is $\frac{1}{2}$. In the second half of the lectures, we turn to the Ising model. Our goal is also to prove that there is a phase transition in any dimension larger than 2. We achieve this by coupling the Ising model with the random cluster model, comparing the latter with percolation, and using our previous results on percolation.

Most of the material presented here is borrowed from two books by Geoffrey Grimmett [1, 3]. I have also used personal notes [2] from a series of lectures on Interacting Particle Systems delivered by G. Grimmett at Cambridge University in 2001.

Giving these lectures has been a great pleasure and I would like to thank the members of the Probability department at Tsinghua University, in particular Professor Wen, for their kind invitation. I am also grateful to all the students for the interest and tenacity that they have shown.

Part I

Percolation

1 Binomial Galton-Watson trees

1.1 Galton-Watson processes

Let μ be a probability measure on $\mathbb{N} = \{0, 1, 2, \dots\}$. Let $(Y_{n,i})_{n,i \geq 1}$ be an infinite array of independent identically distributed random variables with distribution μ . We define recursively a sequence of random variables $(Z_n)_{n \geq 0}$ by setting $Z_0 = 1$ and, for all $n \geq 1$,

$$Z_n = \sum_{i=1}^{Z_{n-1}} Y_{n,i}.$$

By definition, the sequence $(Z_n)_{n \geq 0}$ is a *Galton-Watson process* with *offspring distribution* μ . It is a model for the evolution of a population, according to the following rule: at each unit of time, each living individual gives birth, independently of all other individuals, to a progeny whose size is distributed according to μ , and then dies instantly.

The process $(Z_n)_{n \geq 0}$ is a Markov chain on \mathbb{N} , for which 0 is an absorbing state. The event $\{\exists n \geq 1 : Z_n = 0\}$ is called the extinction of the population. In order to study this event, we introduce some quantities related to μ . Firstly, the generating function of μ is the function $\phi(s) = \sum_{k=0}^{+\infty} s^k \mu(\{k\})$. It is defined at least for $s \in [-1, 1]$. It is continuous on this interval, and it is non-decreasing and convex on $[0, 1]$. The (possibly infinite) left derivative of ϕ at 1 is the mean of μ , denoted by $m = \sum_{k=0}^{+\infty} k \mu(\{k\}) \in [0, +\infty]$.

The first fundamental result of the theory of Galton-Watson processes is the following (see [4], Chapter 0).

Theorem 1.1 *Let $(Z_n)_{n \geq 0}$ be a Galton-Watson process with offspring distribution μ . Then the probability of extinction of the process Z is the smallest solution of the equation $\phi(s) = s$ in the interval $[0, 1]$. In particular, the population gets extinct with probability 1 if and only if $m \leq 1$ and $\mu \neq \delta_1$.*

Let us compute the probability of extinction in a very simple case. Let us choose and a real $p \in [0, 1]$, and let μ_p denote the binomial distribution with parameters 2 and p : for each $k \in \{0, 1, 2\}$, $\mu_p(\{k\}) = \binom{2}{k} p^k (1-p)^{2-k}$. The mean of μ_p is $m_p = 2p$. The generating function ϕ_p of μ_p is $\phi_p(s) = (1-p+sp)^2$. The smallest solution of $\phi(s) = s$ on $[0, 1]$ is

$$\mathbb{P}(\exists n \geq 1 : Z_n = 0) = \begin{cases} 1 & \text{if } p \leq \frac{1}{2} \\ \left(\frac{1-p}{p}\right)^2 & \text{if } p \geq \frac{1}{2}. \end{cases}$$

This function is represented on Figure 1. If we vary p around the value $\frac{1}{2}$, we see a dramatic change in the behaviour of the process. For $p \leq \frac{1}{2}$, the process gets extinct

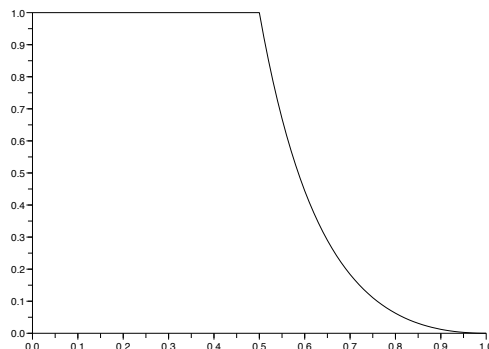


Figure 1: The probability of extinction of a Galton-Watson process with binomial offspring distribution $\mathcal{B}(2, p)$, in function of p . The critical value is $p = \frac{1}{2}$.

with probability 1. For $p > \frac{1}{2}$, there is a positive probability that the process survives forever. Moreover, the probability of extinction is not differentiable with respect to p at the critical value $p = \frac{1}{2}$.

This phenomenon is an example of a *phase transition* : a continuous change of a parameter of the system affects in a non-regular way its global behaviour.

1.2 A tree with branches removed at random

Consider the infinite rooted binary tree T , the graph a part of which is depicted in the left part of Figure 2. There is one special vertex in this tree, the lowest on the picture, with only two neighbours. It is called the *root* and denoted by o . All other vertices have three neighbours. Each vertex has a certain *height*, which is the number of edges which one must traverse to reach o from this vertex. For all $n \geq 0$, there are exactly 2^n vertices which have height n .

Now let us choose $p \in [0, 1]$. Imagine that we go successively over all edges of the tree, and choose, independently for each edge, to keep it with probability p or to remove it with probability $1 - p$. We get a subgraph K of the original tree. An example of this construction is given in the middle part of Figure 2. Then, consider the connected component of the root in K , that is, the set of vertices which are joined to o by edges which belong to P . This connected component, which we denote by $C(o)$, is a subtree of the original tree. Finally, let us define the process $(W_n)_{n \geq 0}$ by declaring that for all $n \geq 0$, W_n is the number of vertices at height n in $C(o)$.

Then the process W is a Galton-Watson process with offspring distribution $\mathcal{B}(2, p)$. Let us explain this fact, rather than give a rigorous proof. The crucial point is that each edge of T is kept or removed independently of all other edges. Consider a vertex v of height n . It is joined in T to two vertices at height $n + 1$, say w_1 and w_2 . Conditionally to the event that v is joined to the root in K , each of w_1 and w_2 is joined to the root, independently

and independently of all other vertices, with probability p . Hence, conditional on the presence of a vertex at height n in $C(o)$, there is a random number of vertex at height $n + 1$, with distribution $\mathcal{B}(2, p)$.

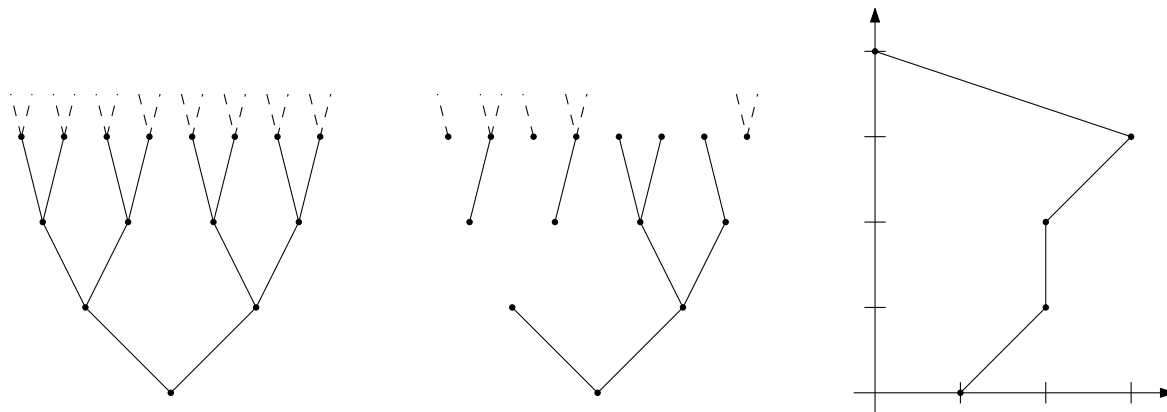


Figure 2: The infinite binary tree, a subtree obtained by removing some edges at random, and the profile of the connected component of the root.

According to the discussion of Section 1.1, the subtree C of T is almost surely finite if $p \leq \frac{1}{2}$ and infinite with positive probability otherwise. This is the prototype of a situation that we are going to study in more detail in these lectures.

The process of *percolation* on a graph consists in removing some edges of this graph chosen at random, independently of each other, with some probability $1 - p$. The result of this operation is a random subgraph of the original graph and a typical question that one can ask is: what is the probability that the connected component in this random subgraph of a given vertex is infinite? We will see that it is usually much harder to answer this question than in the case of a regular tree.

2 The percolation process

2.1 Introduction

Percolation was first considered by Broadbent and Hammersley in 1957, as a model for the diffusion of water in a porous medium, like for instance a porous rock. They imagined that the rock was traversed by many very small channels which could be either open or closed, allowing water to flow through them or not. The question they asked was: how deep can water permeate such a rock? They represented the rock by the network of all the channels through which the water could potentially flow, and then they made the assumption that each channel was open with some fixed probability $p \in [0, 1]$, independently of all other channels. The sub-graph formed by all open channels, which is a random subgraph of the initial graph, is the main object of interest in percolation theory. It turns out that, in

many cases, its large-scale properties depend on the value of p : there are phase transitions in the percolation model, and this is what we are going to study.

Figure 3 shows simulations of the percolation process in a box of a square lattice in the plane.

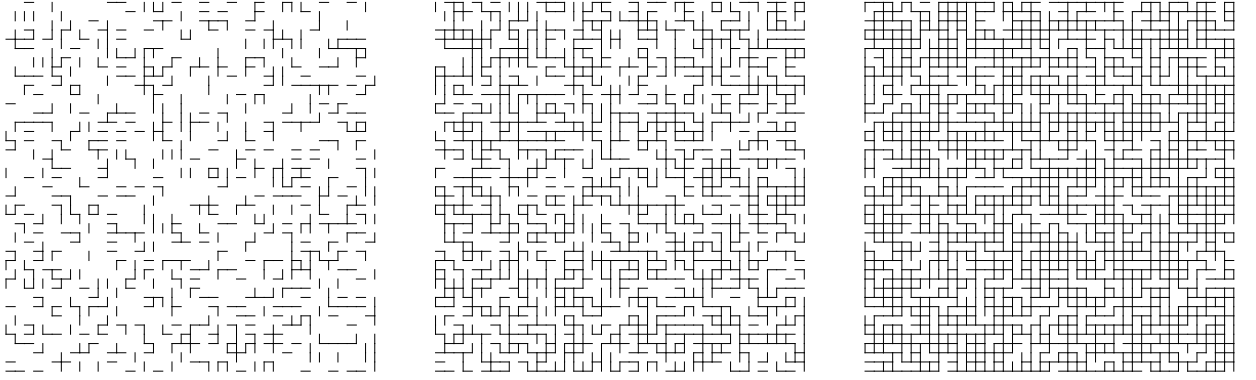


Figure 3: Three simulations of the percolation process in a 40×40 box of the planar square lattice. Only open edges are represented. From left to right, the probability p that each edge is open is equal to 0.25, 0.5 and 0.75.

2.2 Graphs and edge configurations

Percolation is a model of random subgraphs of a graph and we begin by introducing some purely deterministic notions about graphs. Among the many possible definitions of a graph, we choose one where edges are not oriented, and neither multiple edges nor loops are allowed. For all set V , we denote by $\mathcal{P}_2(V)$ the set of pairs of elements of V .

Definition 2.1 *A graph is a pair of countable sets $G = (V, E)$ such that E is a subset of $\mathcal{P}_2(V)$. The elements of V are called vertices and the elements of E edges. A graph is finite if V (hence E) is a finite set.*

An edge is by definition a pair $e = \{x, y\}$ with $x, y \in V$. The edge $\{x, y\}$ is said to join x to y . We write sometimes $x \sim y$ to indicate that the edge $\{x, y\}$ belongs to E .

A graph which we will meet very often is the following. Choose an integer $d \geq 1$. Set $V = \mathbb{Z}^d$. Then, join by an edge any two points which are nearest neighbours, that is, set $E = \{\{x, y\} \in \mathcal{P}_2(\mathbb{Z}^d) : \|x - y\| = 1\}$, where $\|\cdot\|$ denotes the usual Euclidean norm. The graph $\mathbb{L}^d = (\mathbb{Z}^d, \mathbb{E}^d)$ is called the d -dimensional cubic lattice.

Our main object of study will be edge configurations.

Definition 2.2 *Let $G = (V, E)$ be a graph. By an edge configuration on G , we mean a subset of E .*

An edge configuration can be represented either as a subset $K \subset E$, or as a function from E to $\{0, 1\}$, that we denote by $\omega(K)$ or simply ω and which is defined by

$$\forall e \in E, \quad \omega(e) = \begin{cases} 1 & \text{if } e \in K \\ 0 & \text{if } e \notin K. \end{cases}$$

The correspondence between K and ω is bijective and, when a configuration is described by a function $\omega : E \rightarrow \{0, 1\}$, we will use the notation $K(\omega) = \{e \in E : \omega(e) = 1\}$ for the corresponding subset of E .

When we consider a specific edge configuration K , we say that the edges of K are *open* and the edges of $E \setminus K$ are *closed*.

Let $G = (V, E)$ be a graph. A *path* in G is an alternating sequence $(x_0, e_1, x_1, \dots, e_n, x_n)$ of vertices (x_0, \dots, x_n) and edges (e_1, \dots, e_n) , which starts and ends with a vertex, and such that for each $i \in \{1, \dots, n\}$, the edge e_i joins x_{i-1} to x_i . A path $(x_0, e_1, x_1, \dots, e_n, x_n)$ is said to join x_0 to x_n .

Definition 2.3 *Let ω be an edge configuration on G . Let x and y be two vertices of G . We say that x and y are connected by a path of open edges if there exists a path $(x = x_0, e_1, x_1, \dots, e_n, x_n = y)$ such that for all $i \in \{1, \dots, n\}$, $\omega(e_i) = 1$. When x and y are connected by a path of open edges, we write $x \leftrightarrow y$.*

Exercise 2.4 Check that, given an edge configuration ω , the relation “to be connected by a path of open edges” is an equivalence relation on V . The equivalence classes are called the *clusters* of ω .

When an edge configuration is given, we denote the cluster of a vertex x by $C(x)$.

2.3 The measurable space of percolation

Let $G = (V, E)$ be a graph. Let $\Omega = \{0, 1\}^E$ be the space of all functions from E to $\{0, 1\}$, in other words the space of all edge configurations on G . In order to define a probability measure on Ω , we must endow it with a σ -field. When G is finite, we take the σ -field of all subsets of Ω . However, when Ω is infinite, we use a smaller σ -field, called the *cylinder σ -field*.

Definition 2.5 *A subset $A \subset \Omega$ is called a cylinder set if there exist a subset $F \subset E$ and a subset B of $\{0, 1\}^F$ such that*

$$A = \{\omega \in \Omega : \omega|_F \in B\}. \tag{1}$$

A cylinder set which satisfies (1) is said to be based on F . If F can be chosen as a finite set, then A is called a finite cylinder set.

Consider $F \subset E$. The set of all cylinder sets based on F is denoted by \mathcal{F}_F .

Definition 2.6 *The cylinder σ -field is the smallest σ -field on Ω which contains all finite cylinder sets. We denote it by \mathcal{F} .*

Loosely speaking, the events that we allow ourselves to consider are those whose occurrence depends only on the configuration of a finite set of edges, and limits (countable set operations) thereof.

Exercise 2.7 *Let x and y be two vertices of \mathbb{L}^d . Prove that the event $\{x \leftrightarrow y\}$ is not a finite cylinder set. On the other hand, prove that the event that x and y are joined by an open path of length at most 100 is a finite cylinder set. Prove that $\{x \leftrightarrow y\}$ belongs to \mathcal{F} .*

2.4 The percolation measure

Let $G = (V, E)$ be a graph. Choose a real number $p \in [0, 1]$. Let us call μ_p the Bernoulli measure on $\{0, 1\}$ such that $\mu_p(\{1\}) = p$ and $\mu_p(\{0\}) = 1 - p$. We define the probability measure \mathbb{P}_p on $(\Omega = \prod_{e \in E} \{0, 1\}, \mathcal{F})$ as the product measure

$$\mathbb{P}_p = \bigotimes_{e \in E} \mu_p.$$

Since the σ -field \mathcal{F} is generated by finite cylinder sets, the measure \mathbb{P}_p is fully characterized by its value on these sets. If F is a finite subset of E and π is an element of $\{0, 1\}^F$, then the set A defined by (1) satisfies

$$\mathbb{P}_p(A) = \mathbb{P}_p(\forall e \in F, \omega(e) = \pi(e)) = p^{\#\{e \in F: \pi(e)=1\}} (1-p)^{\#\{e \in F: \pi(e)=0\}}.$$

There is no other rule to make computations for the percolation process, and it can be expressed by saying that all edges are open with probability p , independently of each other.

Recall that, for $x \in V$, we denote by $C(x)$ the cluster of x . We denote also by $|C(x)|$ the cardinal of $C(x)$.

Exercise 2.8 *Let x be a vertex of \mathbb{L}^d . Prove that $\mathbb{P}_p(C(x) = \{x\}) = (1-p)^{2d}$. Prove that $\mathbb{P}_p(|C(x)| = 2) = 2dp(1-p)^{4d-2}$.*

An event which we will study in detail is the following.

Lemma 2.9 *Let $\omega \in \Omega$ be an edge configuration on \mathbb{L}^d . Let 0 denote the origin of \mathbb{Z}^d . The following two properties of ω are equivalent.*

1. *The cluster of 0 contains infinitely many vertices : $|C(0)| = +\infty$.*
2. *For all $n \geq 1$, there exists a vertex x such that $\|x\| \geq n$ and $0 \leftrightarrow x$.*

Proof. The subset $C(0)$ of \mathbb{Z}^d contains 0, and it is infinite if and only if it contains vertices with arbitrary large norm. \square

We write $0 \leftrightarrow \infty$ when one of these equivalent properties is satisfied, and sometimes we say that *percolation* occurs. We will detect the existence of a phase transition for the percolation on certain graphs by studying the probability that percolation occurs. On \mathbb{L}^d , we give a name to this probability.

Definition 2.10 Consider the graph \mathbb{L}^d . For all $p \in [0, 1]$, we define the (*d*-dimensional) percolation probability by

$$\theta_d(p) = \mathbb{P}_p(0 \leftrightarrow \infty).$$

2.5 Order and increasing events

A fundamental feature of the configuration space of percolation is that it carries a natural *partial order*.

Definition 2.11 Let $G = (V, E)$ be a graph. Consider ω_1 and ω_2 in $\Omega = \{0, 1\}^E$. It is equivalent to say that $K(\omega_1) \subset K(\omega_2)$ or to say that for all $e \in E$, $\omega_1(e) \leq \omega_2(e)$. When these properties hold, we write $\omega_1 \leq \omega_2$.

In English, $\omega_1 \leq \omega_2$ means that all edges which are open for ω_1 are also open for ω_2 . This is only a partial order on Ω in the sense that, in general, given two configurations ω_1 and ω_2 , it is neither true that $\omega_1 \leq \omega_2$ nor that $\omega_2 \leq \omega_1$. Nevertheless, it makes sense to say that a real-valued function on the configuration space is increasing.

Definition 2.12 1. A function $f : \Omega \rightarrow \mathbb{R}$ is increasing if, for all $\omega_1, \omega_2 \in \Omega$ such that $\omega_1 \leq \omega_2$, the inequality $f(\omega_1) \leq f(\omega_2)$ holds.
2. An event $A \subset \Omega$ is increasing if, for all $\omega_1, \omega_2 \in \Omega$ such that $\omega_1 \leq \omega_2$,

$$\omega_1 \in A \Rightarrow \omega_2 \in A.$$

A function f (resp. an event A) is said to be decreasing if $-f$ (resp. $\Omega \setminus A$) is increasing.

The relation between increasing functions and increasing events is the following: an event A is increasing if and only if the indicator function of A is increasing.

An example of an increasing function is $|C(x)|$ for some vertex x . The prototype of an increasing event is $\{x \leftrightarrow y\}$, for $x, y \in V$. Also, for example, if x is a vertex, the event $\{|C(x)| \leq 100\}$ is decreasing.

Exercise 2.13 Prove that the event $\{x \leftrightarrow y\}$ is an increasing event. Find an event (for instance a finite cylinder set) which is neither increasing nor decreasing.

2.6 Approximation results

Often in percolation theory, one starts by considering events which depend only on finitely many edges, that is, events which belong to \mathcal{F}_F for some finite $F \subset E$, and then one uses an argument of approximation. When increasing events are involved, it is useful to know that they can be approximated by *increasing* events which depend only on finitely many edges. In this paragraph, we collect some useful results.

Firstly, we establish a formula for the conditional expectation of a function $f : \Omega \rightarrow \mathbb{R}$ with respect to a σ -field \mathcal{F}_F . So, let F be a subset of E , which needs not be finite. Let $f : \Omega \rightarrow \mathbb{R}$ be a function which is integrable with respect to \mathbb{P}_p . We want to compute the conditional expectation $\mathbb{E}_p[f|\mathcal{F}_F]$, where \mathbb{E}_p denotes the expectation with respect to \mathbb{P}_p . Let us decompose Ω according to the partition $E = F \cup (E \setminus F)$ of E :

$$\Omega = \{0, 1\}^E = \{0, 1\}^F \times \{0, 1\}^{E \setminus F}.$$

For all $\omega \in \Omega$, we denote by $(\omega_F, \omega_{E \setminus F})$ the components of ω with respect to this decomposition. The measure \mathbb{P}_p decomposes into two independent parts as

$$\mathbb{P}_p = \mathbb{P}_p^F \otimes \mathbb{P}_p^{E \setminus F} = \bigotimes_{e \in E} \mu_p \otimes \bigotimes_{e \in E \setminus F} \mu_p.$$

Proposition 2.14 *Let $f : \Omega \rightarrow \mathbb{R}$ be a \mathbb{P}_p -integrable function. Let F be a subset of E . Then*

$$\mathbb{E}_p[f|\mathcal{F}_F](\eta) = \int_{\{0, 1\}^{E \setminus F}} f(\eta_F, \omega_{E \setminus F}) \mathbb{P}_p^{E \setminus F}(d\omega_{E \setminus F}).$$

The proof is given under the form of an exercise.

Proof. For all $\eta \in \Omega$, define

$$h(\eta) = \int_{\{0, 1\}^{E \setminus F}} f(\eta_F, \omega_{E \setminus F}) \mathbb{P}_p^{E \setminus F}(d\omega_{E \setminus F}).$$

Let us assume that the function f is bounded.

1. Check that $h : \Omega \rightarrow \mathbb{R}$ is well defined and measurable with respect to \mathcal{F}_F .
2. Let $g : \{0, 1\}^E \rightarrow \mathbb{R}$ be a function. Prove that $\mathbb{E}_p[fg] = \mathbb{E}_p[hg]$. Conclude that $\mathbb{E}_p[f|\mathcal{F}_F] = h$ almost surely.
3. Prove that h is well defined even without the assumption that f is bounded. [Hint: use Fubini's theorem.] Conclude. \square

Corollary 2.15 *Let $f : \Omega \rightarrow \mathbb{R}$ be an increasing \mathbb{P}_p -integrable function. Let F be a subset of E . Then $\mathbb{E}_p[f|\mathcal{F}_F]$ is an increasing function.*

Then, the typical approximation argument is the following. We say that a function $f : \Omega \rightarrow \mathbb{R}$ is a *finite cylinder function* if there exists a finite set $F \subset E$ such that f is measurable with respect to \mathcal{F}_F .

Proposition 2.16 *Let $G = (V, E)$ be a graph. Choose $p \in [0, 1]$. Let $f : \Omega \rightarrow \mathbb{R}$ be an increasing function. Assume that $f \in L^r(\mathbb{P}_p)$ for some $r \geq 1$. Then there exists a sequence $(f_n)_{n \geq 0}$ of increasing finite cylinder functions belonging to $L^r(\mathbb{P}_p)$, which converges to f \mathbb{P}_p -almost surely and in $L^r(\mathbb{P}_p)$.*

Proof. If G is a finite graph, there is nothing to prove. Let us assume that G is infinite and recall that, by definition of a graph, the set E is countable. Let $(F_n)_{n \geq 0}$ be a sequence of finite subsets of E such that, for all $n \leq m$, one has $F_n \subset F_m$, and such that $\bigcup_{n \geq 0} F_n = E$. For all $n \geq 0$, set $f_n = \mathbb{E}[f | \mathcal{F}_{F_n}]$. By Proposition 2.15, the function f_n is increasing for all $n \geq 0$. It is a finite cylinder function by definition. Moreover, the sequence $(f_n)_{n \geq 0}$ is a martingale under \mathbb{P}_p , which is bounded in L^r . If $r > 1$, this suffices to imply the result, by the classical theorem of convergence of martingales. If $r = 1$, we need to observe that the martingale is, by definition, uniformly integrable. \square

2.7 Coupling the percolation processes

Intuitively, increasing events are more likely to happen when more edges are open. Hence, we expect their probability to increase when one increases the parameter p . This will be our first non-trivial result, and we will deduce it from the existence a nice coupling of the probabilities \mathbb{P}_p .

Let $G = (V, E)$ be a graph. There is a very nice way of realizing the percolation processes on G , associated to all possible values of $p \in [0, 1]$, at once on a single probability space. For this, consider, on a probability space $(\Xi, \mathcal{A}, \mathbb{P}_*)$ (which we will not use outside this section), a family $(U_e)_{e \in E}$ of independent identically distributed random variables with uniform distribution on the segment $[0, 1]$. Then, for each $p \in [0, 1]$ and each $e \in E$, define a random variable $\omega_p(e)$ on $(\Xi, \mathcal{A}, \mathbb{P}_*)$ by setting

$$\omega_p(e) = \begin{cases} 1 & \text{if } U_e \leq p \\ 0 & \text{if } U_e > p. \end{cases}$$

Then, for all $p \in [0, 1]$, the random vector $\omega_p = (\omega_p(e))_{e \in E}$ has the distribution \mathbb{P}_p . Indeed, the random variables $(\omega_p(e))_{e \in E}$ are independent, identically distributed, with distribution $p\delta_1 + (1 - p)\delta_0$.

The fact that ω_p has the distribution \mathbb{P}_p for all p can be phrased as follows: for all $p \in [0, 1]$ and all event $A \in \mathcal{F}$, the expectation under \mathbb{P}_* of the random variable $\mathbb{1}_A(\omega_p)$ is

$$\mathbb{E}_{\mathbb{P}_*}[\mathbb{1}_A(\omega_p)] = \mathbb{P}_p(A). \quad (2)$$

The point which makes this coupling useful is that for all $p_1 \leq p_2$ in $[0, 1]$, one has \mathbb{P}_* -almost surely $\omega_{p_1} \leq \omega_{p_2}$.

Proposition 2.17 *Let $G = (V, E)$ be a graph. Let $A \in \mathcal{F}$ be an increasing event. Then the function $p \mapsto \mathbb{P}_p(A)$ is non-decreasing on $[0, 1]$.*

Proof. Choose $p_1 \leq p_2$. Since $\omega_{p_1} \leq \omega_{p_2}$ and A is increasing, $\mathbb{1}_A(\omega_{p_1}) \leq \mathbb{1}_A(\omega_{p_2})$, with \mathbb{P}_* -probability 1. Taking the expectation with respect to \mathbb{P}_* , we find, thanks to (2), $\mathbb{P}_{p_1}(A) \leq \mathbb{P}_{p_2}(A)$. \square

The proof of Proposition 2.17 is very short (and actually perhaps a bit obscure) once the coupling of the percolation processes has been defined. But it is not so easy to prove this result without introducing this clever coupling.

Corollary 2.18 For all $d \geq 1$, the percolation probability $\theta_d(p)$ is a non-decreasing function of p . Moreover, $\theta_d(0) = 0$ and $\theta_d(1) = 1$.

Proof. By definition, $\theta_d(p)$ is the probability under \mathbb{P}_p of the event $\{0 \leftrightarrow \infty\}$, which is increasing. The second assertion is straightforward. \square

2.8 The phase transition phenomenon

The example of the binary tree that we have studied as an introduction was a genuine example of a percolation process, on a specific graph which is the infinite rooted binary tree, and the probability of extinction that we have computed there is what we would now call $1 - \theta(p)$. We have thus proved, in this example, that $\theta(p)$ may not be a regular function of p , indeed a non-differentiable one. We will focus on a coarser property of θ , which we have also checked in our example, which is the fact that $\theta(p) = 0$ for some positive values of p and $\theta(p) > 0$ for other positive values of p . Since θ is non-decreasing, going from 0 to 1 on $[0, 1]$, this prevents it at least from being analytic.

Definition 2.19 Consider the graph \mathbb{L}^d . The critical probability is the real $p_c(d) \in [0, 1]$ defined by

$$p_c(d) = \sup\{p \in [0, 1] : \theta_d(p) = 0\}.$$

The case of the one-dimensional lattice \mathbb{L}^1 is exceptional in that it satisfies $p_c(1) = 1$.

Exercise 2.20 Consider the percolation process on \mathbb{L}^1 . Choose $p < 1$. By considering the intersection of $C(0)$ with $\mathbb{Z}_+^* = \{1, 2, 3, \dots\}$ and $\mathbb{Z}_-^* = \{\dots, -3, -2, -1\}$, prove that $|C(0)| - 1$ has the distribution of the sum of two independent geometric random variables. Conclude that $\mathbb{P}_p(0 \leftrightarrow \infty) = 0$. Prove finally that $p_c(1) = 1$.

The first very important result on the percolation process on \mathbb{L}^d states that there is a non-trivial phase transition whenever $d \geq 2$.

Theorem 2.21 (Phase transition for the percolation on \mathbb{L}^d) Assume that $d \geq 2$. Then $0 < p_c(d) < 1$.

It is extremely hard in general to compute $p_c(d)$, or even to find good explicit bounds. H. Kesten proved in 1980 that $p_c(2) = \frac{1}{2}$, using to this end partial results which had been collected for more than twenty years. At the end of our study of percolation, we will prove a part of this statement.

Fortunately, it is considerably simpler to prove Theorem 2.21. We will do it by considering *self-avoiding paths*. By definition, a self-avoiding path is a path which does not visit twice any vertex, that is, a path $(x_0, e_1, x_1, \dots, e_n, x_n)$ such that all the vertices x_0, \dots, x_n are distinct. This implies of course that the edges e_1, \dots, e_n are also distinct. Let us denote by $\kappa_n(d)$ the number of distinct self-avoiding paths of length n which start from the origin in \mathbb{L}^d .

Lemma 2.22 1. For all $n \geq 1$, $\kappa_n(d) \leq 2d(2d-1)^{n-1}$.

2. Let $\omega \in \Omega$ be a configuration such that $|C(0)| = +\infty$. Then for all $n \geq 1$, there exists an open self-avoiding path of length n starting at 0.

Proof. 1. The first step of a self-avoiding path can be any of the $2d$ edges adjacent to its starting point. Then, at each step, it can certainly not leave its current position through the edge that it has used to get there. Thus, there are at most $2d-1$ available edges.

2. Since $|C(0)| = +\infty$, $C(0)$ contains at least one vertex x such that $\|x\| \geq n$. An open path of minimal length joining 0 to x is necessarily self-avoiding and of length larger than n . By stopping such a path at its n -th step, we get the desired path. \square

Proof of Theorem 2.21 ($p_c(d) > 0$) – We claim that $p_c(d) \geq \frac{1}{2d-1}$. Indeed, choose $p < \frac{1}{2d-1}$. Let us estimate $\theta_d(p)$. Let us choose an integer $n \geq 1$. By the second assertion of the preceding lemma,

$$\theta_d(p) = \mathbb{P}_p(0 \leftrightarrow \infty) \leq \mathbb{P}_p(C(0) \text{ contains a self-avoiding path of length } n).$$

Let N_n denote the number of self-avoiding paths of length n contained in $C(0)$. We have

$$\theta_d(p) \leq \mathbb{P}_p(N_n \geq 1) \leq \mathbb{E}_p(N_n) = p^n \kappa_n(d) \leq 2d((2d-1)p)^n.$$

Since $(2d-1)p < 1$, the last quantity tends to 0 as n tends to infinity. Hence, $\theta_d(p) = 0$. \square

In order to prove that $p_c(d) < 1$, we start by reducing the problem to the 2-dimensional case.

Lemma 2.23 If $d \leq d'$, then $p_c(d') \leq p_c(d)$.

In particular, it suffices to prove the inequality $p_c(2) < 1$ in order to prove that $p_c(d) < 1$ for all $d \geq 2$.

Proof. Choose $d \leq d'$. The lattice \mathbb{L}^d can be viewed as a sub-lattice of the lattice $\mathbb{L}^{d'}$, for instance by identifying $(x_1, \dots, x_d) \in \mathbb{Z}^d$ with $(x_1, \dots, x_d, 0, \dots, 0) \in \mathbb{Z}^{d'}$. The percolation process on $\mathbb{L}^{d'}$ restricted to \mathbb{L}^d is the percolation process on \mathbb{L}^d . In other words, we have a coupling of the percolation processes on \mathbb{L}^d and $\mathbb{L}^{d'}$. Now, in any configuration, the cluster of the origin in \mathbb{L}^d is contained in the cluster of the origin in $\mathbb{L}^{d'}$. Hence, $\theta_d(p) = P_p(0 \leftrightarrow \infty \text{ in } \mathbb{L}^d) \leq P_p(0 \leftrightarrow \infty \text{ in } \mathbb{L}^{d'}) \leq \theta_{d'}(p)$. The result follows. \square

There remains to prove that $p_c(2) < 1$. For this, we introduce the dual lattice \mathbb{L}_*^2 . This lattice is built by putting one vertex in the middle of each face of \mathbb{L}^2 and by joining any two vertices which are in the centres of two faces of \mathbb{L}^2 which share a bounding edge (see the left part of Figure 4). The lattice \mathbb{L}_*^2 is isomorphic to \mathbb{L}^2 , it is in fact simply the lattice \mathbb{L}^2 translated by the vector $(\frac{1}{2}, \frac{1}{2})$.

Each edge of \mathbb{L}^2 crosses exactly one edge of \mathbb{L}_*^2 . Hence, there is a natural way to deduce a configuration on \mathbb{L}_*^2 from a configuration on \mathbb{L}^2 : we declare simply that the open edges

of \mathbb{L}_*^2 are those which cross open edges of \mathbb{L}^2 . We denote by ω_* the configuration on \mathbb{L}_*^2 deduced in this way from a configuration ω on \mathbb{L}^2 and we call it the *dual configuration*. It is the interplay between configurations on the lattice and its dual which make the 2-dimensional percolation somewhat easier to study than the general d -dimensional process, and in particular which allow one to prove that $p_c(2) = \frac{1}{2}$.

Let us call a path $(x_0, e_1, x_1, \dots, x_{n-1}, e_n, x_n)$ a *self-avoiding loop* if $(x_0, e_1, x_1, \dots, x_{n-1})$ is a self-avoiding path and $x_n = x_0$. The range of a self-avoiding loop is a Jordan curve, hence it separates the plane in two connected components, one of which is bounded and is called the interior of the loop, and the other which is unbounded and called the exterior. The following result is intuitively obvious, but to actually prove it rigorously takes quite of a lot of effort. We won't do it in these notes.

Proposition 2.24 *Let $\omega \in \Omega$ be a configuration on \mathbb{L}^2 . Assume that $C(0)$ is finite. Then there exists a closed self-avoiding loop in \mathbb{L}_*^2 whose interior contains $C(0)$.*

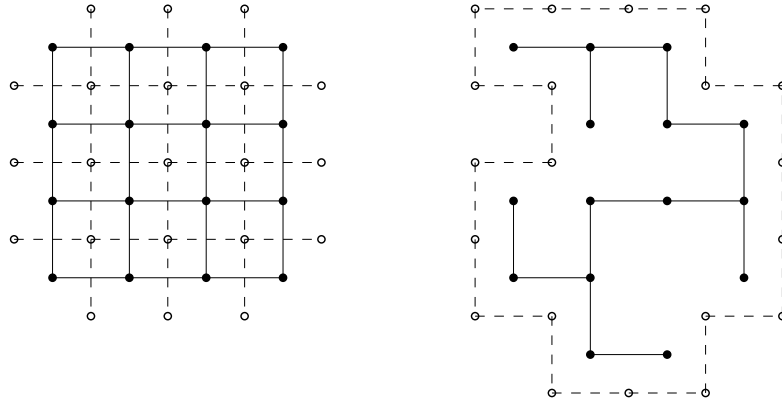


Figure 4: The dual graph \mathbb{L}_*^2 is isomorphic to \mathbb{L}^2 . Any finite open cluster in \mathbb{L}^2 is surrounded by a closed self-avoiding loop in the dual lattice.

The right part of Figure 4 illustrates this fact. Let us estimate the number of self-avoiding loops of a given length which surround 0.

Lemma 2.25 *Let $n \geq 4$ be an integer. The number of self-avoiding loops of length n in \mathbb{L}_*^2 which surround the origin of \mathbb{L}^2 is smaller than $n\kappa_{n-1}(2)$.*

Proof. Let γ be a self-avoiding loop of length n of \mathbb{L}_*^2 which surrounds the origin. It must visit a vertex of the form $(k + \frac{1}{2}, \frac{1}{2})$ for some $k \in \{0, \dots, n-1\}$. Let x be the rightmost vertex of this form visited by γ . If we go along γ , starting at x and turning anticlockwise around 0, during $n-1$ steps, we find a self-avoiding path of length $n-1$. Since γ is determined by x and the self-avoiding path, the number of self-avoiding loops of length n is smaller than n times the number of self-avoiding paths of length $n-1$. \square

Proof of Theorem 2.21 ($p_c(d) < 1$) – By Proposition 2.24,

$$\begin{aligned} 1 - \theta_2(p) &= \mathbb{P}_p(0 \text{ is surrounded by a closed dual self-avoiding loop}) \\ &\leq \mathbb{E}_p[\text{number of closed dual self-avoiding loop surrounding } 0] \\ &\leq \sum_{n=4}^{+\infty} (1-p)^n n \kappa_{n-1}(2). \end{aligned}$$

Since $\kappa_{n-1}(2) \leq 4^{n-1}$, the sum of the last series is strictly smaller than 1 if p is close enough to 1. Hence, $\theta_2(p) > 0$ for p close enough to 1. \square

It is not difficult to check that the argument above implies that $p_c(2) \leq \frac{7}{8}$. Finally, we have proved that $p_c(2) \in [\frac{1}{3}, \frac{7}{8}]$.

3 Correlation inequalities and stochastic ordering

3.1 The FKG inequality

Increasing events are those which, intuitively, tend to happen when many edges are open. It is thus plausible that any two increasing events are positively correlated. This is the content of the FKG inequality, which is named after Fortuin, Kasteleyn and Ginibre.

Proposition 3.1 (FKG inequality) *Consider the percolation process on \mathbb{L}^d with some parameter $p \in [0, 1]$.*

1. *For all increasing events A, B , one has the inequality*

$$\mathbb{P}_p(A \cap B) \geq \mathbb{P}_p(A)\mathbb{P}_p(B).$$

2. *For all increasing functions f, g such that $\mathbb{E}_p[f^2] < \infty$ and $\mathbb{E}_p[g^2] < \infty$, one has the inequality*

$$\mathbb{E}_p[fg] \geq \mathbb{E}_p[f]\mathbb{E}_p[g].$$

Provided $\mathbb{P}_p(A) > 0$, the first assertion can be rewritten as

$$\mathbb{P}_p(B|A) \geq \mathbb{P}_p(B),$$

which expresses more obviously the fact that any two increasing events are positively correlated.

Proof. The second formulation of the FKG inequality applied to the indicator functions of increasing events yields immediately the first formulation. Thus, we focus on the functional version.

We start by proving the result when f and g are finite cylinder functions, by induction on the number of edges on which they depend.

Let us assume that f and g both depend on only one edge $e \in \mathbb{E}^d$, that is, that they are measurable with respect to $\mathcal{F}_{\{e\}}$. Let ω_o denote a configuration such that $\omega_o(e) = 1$ and ω_c a configuration such that $\omega_c(e) = 0$. The assumption that f and g are increasing means that $f(\omega_o) \geq f(\omega_c)$ and $g(\omega_o) \geq g(\omega_c)$. Let us first assume that $f(\omega_c) = g(\omega_c) = 0$. Then

$$\mathbb{E}_p[fg] = pf(\omega_o)g(\omega_o) \geq p^2f(\omega_o)g(\omega_o) = \mathbb{E}_p[f]\mathbb{E}_p[g].$$

Replacing f by $f - f(\omega_c)$ and g by $g - g(\omega_c)$ allows us to remove the last assumption.

Let us now choose $n \geq 2$ and assume that the result has been proved for any pair of functions f and g which depend on at most $n - 1$ edges. Let f and g be two increasing functions which depend only on n edges e_1, \dots, e_n . We proceed by conditioning the expectations with respect to e_1, \dots, e_{n-1} : indeed, $\mathbb{E}_p[f|e_1, \dots, e_{n-1}]$ and $\mathbb{E}_p[g|e_1, \dots, e_{n-1}]$ are two functions which depend only on e_1, \dots, e_{n-1} . From the explicit formula

$$\mathbb{E}_p[f|e_1, \dots, e_{n-1}](\varepsilon_1, \dots, \varepsilon_{n-1}) = pf(\varepsilon_1, \dots, \varepsilon_{n-1}, 1) + (1 - p)f(\varepsilon_1, \dots, \varepsilon_{n-1}, 0), \quad (3)$$

valid for all $\varepsilon_1, \dots, \varepsilon_{n-1} \in \{0, 1\}$, follows immediately the fact that $\mathbb{E}_p[f|e_1, \dots, e_{n-1}]$ is an increasing function. Thus, by induction, the FKG inequality applies and gives

$$\mathbb{E}_p[\mathbb{E}_p[f|e_1, \dots, e_{n-1}]\mathbb{E}_p[g|e_1, \dots, e_{n-1}]] \geq \mathbb{E}_p[\mathbb{E}_p[f|e_1, \dots, e_{n-1}]]\mathbb{E}_p[\mathbb{E}_p[g|e_1, \dots, e_{n-1}]].$$

The right-hand side is equal to $\mathbb{E}_p[f]\mathbb{E}_p[g]$ by the elementary properties of the conditional expectation. To treat the left-hand side, we claim that

$$\mathbb{E}_p[fg|e_1, \dots, e_{n-1}] \geq \mathbb{E}_p[f|e_1, \dots, e_{n-1}]\mathbb{E}_p[g|e_1, \dots, e_{n-1}].$$

It suffices to prove that for all $\varepsilon_1, \dots, \varepsilon_{n-1} \in \{0, 1\}$, the inequality

$$\begin{aligned} \mathbb{E}_p[fg|e_1 = \varepsilon_1, \dots, e_{n-1} = \varepsilon_{n-1}] &\geq \\ \mathbb{E}_p[f|e_1 = \varepsilon_1, \dots, e_{n-1} = \varepsilon_{n-1}]\mathbb{E}_p[g|e_1 = \varepsilon_1, \dots, e_{n-1} = \varepsilon_{n-1}] & \end{aligned}$$

holds. By subtracting a constant to f and g , we may assume that $f(\varepsilon_1, \dots, \varepsilon_{n-1}, 0) = g(\varepsilon_1, \dots, \varepsilon_{n-1}, 0) = 0$. Then the claimed inequality follows from (3) and the same argument that we used to prove the FKG inequality for functions which depend only on one edge (or the more general result stated in Proposition 2.15).

At this point, we have proved the FKG inequality for any pair of increasing functions which depend on a finite number of edges. The general case follows thanks to the approximation argument summarized by Proposition 2.16. \square

Exercise 3.2 *Recall that a decreasing event is an event whose complement is increasing. Study the correlation between two decreasing events, between an increasing and a decreasing event. Extend your results to increasing and decreasing functions.*

3.2 The BK inequality

The BK inequality, discovered by van den Berg and Kesten, provides, in certain situations, an inequality of the same kind as the FKG inequality, but in the opposite direction. The crucial notion is that of *disjoint occurrence* of two events. Intuitively, two increasing events occur disjointly in a given configuration if the set of open edges of this configuration can be partitioned into two subsets such that the presence of the edges of the first (resp. the second) set suffices to guarantee that the first (resp. the second) event is realized.

Definition 3.3 *Let A and B be two events. Let ω be a configuration. The events A and B occur disjointly at ω , and we write $\omega \in A \circ B$ if the following property holds : there exists a partition $K(\omega) = H_A \sqcup H_B$ such that the two configurations ω_A and ω_B determined respectively by $K(\omega_A) = H_A$ and $K(\omega_B) = H_B$ satisfy $\omega_A \in A$ and $\omega_B \in B$.*

Observe that the event $A \circ B$ has been defined without the assumption that A or B is increasing.

Exercise 3.4 *Consider two increasing events A and B . Prove that $A \cap B \supset A \circ B$. Prove that $A \circ B$ is increasing.*

Proposition 3.5 (BK inequality) *Let A and B be two increasing events which depend on finitely many edges. Then*

$$\mathbb{P}_p(A \circ B) \leq \mathbb{P}_p(A)\mathbb{P}_p(B).$$

We will not prove the BK inequality in these notes. A proof can be found in [1], pp. 39–41.

3.3 Stochastic ordering

The proof of the FKG inequality that we have given is simple but it has the drawback of being very special to percolation. In this section, we will give another proof of the FKG inequality, based on the notion of stochastic ordering. It is a much more robust proof, which in fact provides us with a condition for fairly general edge models under which the FKG inequality holds. We will use this condition later when we discuss the random cluster model.

In the present context, the graph structure does not really matter. We are really only considering a set E of edges. We say that a probability measure μ on $\Omega = \{0, 1\}^E$ is *positive* if $\mu(\omega) > 0$ for all $\omega \in \Omega$. The main definition is the following.

Definition 3.6 *Let μ_1 and μ_2 be two probability measures on Ω . We say that μ_2 is stochastically dominated by μ_1 , and we write $\mu_2 \leq_{\text{st}} \mu_1$ if for all increasing event A , one has $\mu_2(A) \leq \mu_1(A)$.*

Equivalently, $\mu_2 \leq_{\text{st}} \mu_1$ if and only if, for all increasing function $f : \Omega \rightarrow \mathbb{R}$, one has $\mu_2(f) \leq \mu_1(f)$.

The partially ordered subset Ω has the important property of being a *lattice*, which means that infima and suprema exist. Concretely, given two configurations ω_1 and ω_2 , we can define two configurations $\omega_1 \vee \omega_2$ and $\omega_1 \wedge \omega_2$, by setting, for all $e \in E$,

$$\omega_1 \vee \omega_2(e) = 1 \text{ if and only if } (\omega_1(e) = 1 \text{ or } \omega_2(e) = 1),$$

$$\omega_1 \wedge \omega_2(e) = 1 \text{ if and only if } (\omega_1(e) = 1 \text{ and } \omega_2(e) = 1).$$

They have the property that $\omega \vee \omega_2$ is the smallest configuration which is greater than both ω_1 and ω_2 , and $\omega_1 \wedge \omega_2$ is the greatest configuration which is smaller than both ω_1 and ω_2 . The main result that we want to prove is the following.

Proposition 3.7 (FKG condition) *Let E be a finite set. Let μ be a positive measure on $\Omega = \{0, 1\}^E$. Assume that for all $\omega_1, \omega_2 \in \Omega$, the inequality*

$$\mu(\omega_1 \vee \omega_2)\mu(\omega_1 \wedge \omega_2) \geq \mu(\omega_1)\mu(\omega_2)$$

holds. Then μ satisfies the FKG inequality: for all $f, g : \Omega \rightarrow \mathbb{R}$ increasing functions,

$$\mu(fg) \geq \mu(f)\mu(g).$$

In order to prove this result, we use a result which is useful on its own and which is known as Holley's inequality.

Proposition 3.8 (Holley's inequality) *Let μ_1 and μ_2 be two positive probability measures on Ω . Assume that for all $\omega_1, \omega_2 \in \Omega$, the following inequality holds:*

$$\mu_1(\omega_1 \vee \omega_2)\mu_2(\omega_1 \wedge \omega_2) \geq \mu_1(\omega_1)\mu_2(\omega_2).$$

Then $\mu_2 \leq_{\text{st}} \mu_1$.

Let us first deduce the FKG condition from Holley's inequality.

Proof of Proposition 3.7 – By adding a constant to the function g if necessary, we may assume that $g > 0$. Define a probability measure μ' on Ω by setting, for all $\omega \in \Omega$

$$\mu'(\omega) = \frac{\mu(\omega)g(\omega)}{\sum_{\pi} \mu(\pi)g(\pi)}.$$

It is a positive measure on Ω . Consider $\omega_1, \omega_2 \in \Omega$. We have

$$\begin{aligned} \mu'(\omega_1 \vee \omega_2)\mu(\omega_1 \wedge \omega_2) &= \frac{1}{\mu(g)}g(\omega_1 \vee \omega_2)\mu(\omega_1 \wedge \omega_2)\mu(\omega_1 \wedge \omega_2) \\ &\geq \frac{1}{\mu(g)}g(\omega_1)\mu(\omega_1 \wedge \omega_2)\mu(\omega_1 \wedge \omega_2) \\ &\geq \frac{1}{\mu(g)}g(\omega_1)\mu(\omega_1)\mu(\omega_2) \\ &= \mu'(\omega_1)\mu(\omega_2). \end{aligned}$$

We have used successively the definition of μ' , the fact that g is increasing, the assumption on μ and finally the definition of μ' again.

By Holley's inequality (Proposition 3.8), this implies that $\mu \leq_{st} \mu'$. Hence, by definition, $\mu'(f) \geq \mu(f)$, that is, $\mu(fg) \geq \mu(f)\mu(g)$, which is the FKG inequality. \square

The proof of Holley's inequality relies on a nice coupling argument and makes use of continuous-time Markov chains.

Proof of Proposition 3.8 – Set $\Omega^* = \{(\omega, \pi) \in \Omega^2 : \omega \geq \pi\}$. The main point is that the assumption on μ_1 and μ_2 allows us to construct a Markov chain on Ω^* whose stationary measure has μ_1 and μ_2 as marginals. Since the Markov chain visits only ordered pairs of configurations, it is then easy to deduce that $\mu_1 \geq_{st} \mu_2$.

In order to define a (continuous-time) Markov chain on Ω^* , we need to specify its generator, which is a matrix G whose entries are indexed by pairs of elements of Ω^* . Let us introduce a notation. Given a configuration $\omega \in \Omega$ and an edge $e \in E$, we define two new configurations ω^e and ω_e by forcing the presence or the absence of e in ω :

$$\forall f \in E, \quad \omega^e(f) = \begin{cases} \omega(f) & \text{if } f \neq e \\ 1 & \text{if } f = e \end{cases} \quad \text{and} \quad \omega_e(f) = \begin{cases} \omega(f) & \text{if } f \neq e \\ 0 & \text{if } f = e \end{cases}.$$

We define now, for all $(\omega, \pi) \in \Omega$ and $e \in E$,

$$\begin{aligned} G((\omega, \pi_e), (\omega^e, \pi^e)) &= 1, \\ G((\omega^e, \pi), (\omega_e, \pi_e)) &= \frac{\mu_1(\omega_e)}{\mu_1(\omega^e)} \\ G((\omega^e, \pi^e), (\omega_e, \pi_e)) &= \frac{\mu_2(\pi_e)}{\mu_2(\pi^e)} - \frac{\mu_1(\omega_e)}{\mu_1(\omega^e)}. \end{aligned}$$

We set $G((\omega, \pi), (\omega', \pi')) = 0$ for all other pairs $(\omega, \pi) \neq (\omega', \pi')$. The diagonal entries of G are finally determined by the fact that all row sums of G must be zero.

The first thing to check is that G is indeed the generator of a Markov chain: it is not clear that $G((\omega^e, \pi^e), (\omega^e, \pi_e))$ is non-negative. In fact, the assumption on μ_1 and μ_2 applied to the configurations ω_e and π^e yields, using the fact that $\pi \leq \omega$,

$$\mu_1(\omega^e)\mu_2(\pi_e) \geq \mu_1(\omega_e)\mu_2(\pi^e),$$

and the non-negativity of the off-diagonal entries of G .

Let $(X_t, Y_t)_{t \geq 0}$ be a Markov chain with generator G . One can check that both components of this process, X and Y , are still Markov chains on Ω . Their respective generators G_X and G_Y can be described as follows. For all $\omega, \pi \in \Omega$ and $e \in E$,

$$\begin{aligned} G_X(\omega_e, \omega^e) &= 1 \quad \text{and} \quad G_X(\omega^e, \omega_e) = \frac{\mu_1(\omega_e)}{\mu_1(\omega^e)}, \\ G_Y(\pi_e, \pi^e) &= 1 \quad \text{and} \quad G_Y(\pi^e, \pi_e) = \frac{\mu_1(\omega_e)}{\mu_1(\omega^e)} + \frac{\mu_2(\pi_e)}{\mu_2(\pi^e)} - \frac{\mu_1(\omega_e)}{\mu_1(\omega^e)} = \frac{\mu_2(\pi_e)}{\mu_2(\pi^e)}. \end{aligned}$$

It is easy to check that for all $\omega_1, \omega_2 \in \Omega$,

$$\mu_1(\omega_1)G_X(\omega_1, \omega_2) = \mu_1(\omega_2)G(\omega_2, \omega_1),$$

and the similar relation for G_Y . This implies that the measure μ_1 (resp. μ_2) is reversible for G_X (resp. G_Y). Since X (resp. Y) is irreducible, μ_1 (resp. μ_2) is its unique invariant probability measure and, as t tends to $+\infty$, X_t (resp. Y_t) converges in distribution to μ_1 (resp. μ_2). Moreover, with probability 1, the inequality $X_t \geq Y_t$ holds for all t by construction.

Let ν denote the invariant measure of (X, Y) . It is some probability measure on Ω^* whose first marginal is μ_1 and whose second marginal is μ_2 . Let $f : \Omega \rightarrow \mathbb{R}$ be increasing. We have

$$\mu_2(f) = \sum_{(\omega, \pi) \in \Omega^*} f(\pi) \nu(\omega, \pi) \leq \sum_{(\omega, \pi) \in \Omega^*} f(\omega) \nu(\omega, \pi) = \mu_1(f).$$

Hence, $\mu_2 \leq_{\text{st}} \mu_1$, which is what we wanted to prove. \square

Exercise 3.9 *Check that the percolation process on a finite graph satisfies the FKG condition. In fact, check that the equality $\mu(\omega_1 \vee \omega_2)\mu(\omega_1 \wedge \omega_2) = \mu(\omega_1)\mu(\omega_2)$ holds. [Hint: check and use the fact that for any two subsets A and B of a finite set, $|A \cup B| + |A \cap B| = |A| + |B|$.]*

4 The cluster of the origin and the number of infinite clusters

In this section, we will discuss the size of the clusters of a typical configuration of the percolation process on \mathbb{L}^d . We already know that this size depends on the value of p : for example, the cluster of the origin has a positive probability of being infinite when $p > p_c(d)$, whereas it is almost surely finite when $p < p_c(d)$. Note that we do not know yet what happens for $p = p_c(d)$. We will prove later that $\theta_2(\frac{1}{2}) = 0$.

We are not only going to consider the cluster of the origin. In fact, we are going to consider the event that there exists an infinite cluster somewhere in the lattice:

$$\{\exists \text{ infinite cluster}\} = \{\exists x \in \mathbb{Z}^d : |C(x)| = +\infty\}.$$

We are also going to discuss the number of distinct infinite clusters present in a typical configuration.

4.1 Existence of an infinite cluster

Let us recall Kolmogorov's 0-1 law, in the context of percolation on \mathbb{L}^d .

Proposition 4.1 (0-1 law) *Let A be an event on the probability space $(\Omega, \mathcal{F}, \mathbb{P}_p)$ of the percolation process on \mathbb{L}^d for some parameter $p \in [0, 1]$. Assume that, for all finite subset $F \subset \mathbb{E}^d$, the event A belongs to $\mathcal{F}_{\mathbb{E}^d \setminus F}$. Then $\mathbb{P}_p(A) \in \{0, 1\}$.*

Proof. The event A is independent of all finite cylinder sets, which generate \mathcal{F} . Hence, it is independent of itself and this implies that its probability is equal to its square, hence to 0 or 1. \square

An event which is independent of any finite set of edges is called an *asymptotic event*.

Exercise 4.2 *Prove that the event $\{\exists \text{ infinite cluster}\}$ is an asymptotic event.*

This allows us to draw the following conclusion.

Proposition 4.3 *If $p < p_c(d)$, there exists almost surely no infinite cluster. If $p > p_c(d)$, there exists almost surely an infinite cluster.*

Proof. Assume that $p > p_c(d)$. Then

$$\mathbb{P}_p(\exists \text{ infinite cluster}) \geq \mathbb{P}_p(|C(0)| = +\infty) = \theta_d(p) > 0,$$

by definition of $p_c(d)$. By the 0-1 law, this implies that $\mathbb{P}_p(\exists \text{ infinite cluster}) = 1$.

Assume now that $p < p_c(d)$. Then $\mathbb{P}_p(|C(0)| = +\infty) = 0$. But the percolation process is invariant by translation and the vertex 0 plays no special role. Hence, $\mathbb{P}_p(|C(x)| = +\infty) = 0$ for all $x \in \mathbb{Z}^d$. On the other hand, the event that there exists an infinite cluster can be written as

$$\{\exists \text{ infinite cluster}\} = \bigcup_{x \in \mathbb{Z}^d} \{|C(x)| = +\infty\}.$$

Hence, it is a countable union of negligible events : it is thus itself negligible, that is, $\mathbb{P}(\exists \text{ infinite cluster}) = 0$, as expected. \square

In two dimensions, we will see that, at the critical probability, there is almost surely no infinite cluster.

4.2 The number of infinite clusters in the supercritical phase

We have proved that in the supercritical phase, that is, for $p > p_c(d)$, there exists almost surely an infinite cluster. We would like to know how many there are. Thus, we define a random variable N which is the number of distinct infinite clusters. This random variable takes its values in $\{0, 1, 2, \dots\} \cup \{+\infty\}$. We have proved that $\mathbb{P}_p(N \geq 1) = 1$ for $p > p_c(d)$.

The first important property of N is that it is almost surely constant: there exists some deterministic element $k(p) \in \{0, 1, 2, \dots\} \cup \{+\infty\}$, which depends on p , such that $\mathbb{P}_p(N = k(p)) = 1$ for all $p \in [0, 1]$. This follows from the ergodic property of the percolation process, which we explain now. For all $w \in \mathbb{Z}^d$, the translation by w acts

on the lattice \mathbb{L}^d and transforms a configuration ω into another configuration which we denote by $\tau_w(\omega)$. We say that a function $f : \Omega \rightarrow \mathbb{R}$ is *translation-invariant* if for all $\omega \in \Omega$ and all $w \in \mathbb{Z}^d$,

$$f(\tau_w(\omega)) = f(\omega).$$

Proposition 4.4 (Ergodic theorem) *Let $f : \Omega \rightarrow \mathbb{R}$ be a translation-invariant function. Then for all $p \in [0, 1]$, the function f is \mathbb{P}_p -almost surely constant.*

Proof. Let $(F_n)_{n \geq 0}$ be an exhausting sequence of \mathbb{L}^d , that is, an increasing sequence of finite subsets whose union is equal to \mathbb{L}^d . Let us assume that f is bounded and, by subtracting a constant if necessary, that $\mathbb{E}_p[f] = 0$. By Proposition 2.16, we know that the sequence $(\mathbb{E}_p[f|\mathcal{F}_{F_n}])_{n \geq 0}$ converges in L^2 to f . Choose $\varepsilon > 0$. There exists $m \geq 0$ such that $\|f - \mathbb{E}_p[f|\mathcal{F}_{F_m}]\|_{L^2} \leq \varepsilon$. Set $g = \mathbb{E}_p[f|\mathcal{F}_{F_m}]$. Since the translations preserve the measure \mathbb{P}_p , it follows, for all $w \in \mathbb{Z}^d$, that

$$\|g - g \circ \tau_w\|_{L^2} \leq \|g - f\|_{L^2} + \|f - f \circ \tau_w\|_{L^2} + \|f \circ \tau_w - g \circ \tau_w\|_{L^2} \leq \varepsilon + 0 + \varepsilon = 2\varepsilon.$$

Now, $g \circ \tau_w$ is measurable with respect to $\mathcal{F}_{F_m - w}$ and, since F_m is finite, it is possible to choose w such that $F_m \cap (F_m - w) = \emptyset$. Let us choose w in this way. Then, g and $g \circ \tau_w$ are independent. On the other hand, the last inequality tells us that they are close in L^2 norm. Thus, they should not be far from being constant. Let us write this rigorously. Since g and $g \circ \tau_w$ are centered, their variance is the square of their L^2 norm. Hence,

$$\text{Var}(g) \leq \text{Var}(g) + \text{Var}(g \circ \tau_w) = \text{Var}(g - g \circ \tau_w) = \|g - g \circ \tau_w\|_{L^2}^2 \leq 4\varepsilon^2.$$

Now, we go back to f :

$$\text{Var}(f) = \|f\|_{L^2}^2 \leq (\|f - g\|_{L^2} + \|g\|_{L^2})^2 \leq 2\|f - g\|_{L^2}^2 + 2\|g\|_{L^2}^2 \leq 10\varepsilon^2.$$

This holds for all $\varepsilon > 0$, so that f is constant. A short argument allows one to remove the assumption that f is bounded. \square

Corollary 4.5 *The random variable N is \mathbb{P}_p -almost surely constant for all $p \in [0, 1]$. We denote its value by $k(p)$.*

We now that $k(p) = 0$ for $p < p_c(d)$ and $k(1) = 1$. The main result of this section is the following.

Theorem 4.6 *For all $p > p_c(d)$, one has $k(p) = 1$.*

In English, this means that in the supercritical phase, the infinite cluster is almost surely unique. We will not prove this theorem completely. We will prove the following partial result.

Proposition 4.7 *For all $p \in [0, 1]$, one has $k(p) \in \{0, 1, +\infty\}$.*

Proof. For all subset $B \subset \mathbb{Z}^d$, let us denote by $N_B(0)$ (resp. $N_B(1)$) the total number of infinite clusters when all edges which join two vertices of B are declared to be closed (resp. open). Thus, for some configuration ω , $N_B(0)(\omega)$ is the number of infinite clusters in the configuration obtained from ω by closing all edges which join two vertices of B .

Consider a finite subset $B \subset \mathbb{Z}^d$. Since there is a positive probability that all edges which join two vertices of B are closed, and since $N = k(p)$ with probability 1, we also have $N_B(0) = k(p)$ with probability 1. Similarly, $N_B(1) = k(p)$ with probability 1. In particular,

$$\mathbb{P}_p(N_B(0) = N_B(1)) = 1.$$

Let $(B_n)_{n \geq 0}$ be an increasing sequence of finite subsets of \mathbb{Z}^d such that $\bigcup_{n \geq 0} B_n = \mathbb{Z}^d$. Assume that $k(p) < +\infty$. Then, almost surely, there exists an integer n such that B_n meets the $k(p)$ infinite clusters. Hence,

$$\mathbb{P}_p\left(\bigcup_{n \geq 0} \{B_n \text{ meets all the } k(p) \text{ infinite clusters}\}\right) = 1.$$

Hence, by the σ -additivity of \mathbb{P}_p , there exists $m \geq 0$ such that

$$\mathbb{P}_p(B_m \text{ meets all the } k(p) \text{ infinite clusters}) > 0.$$

We have thus proved that there exists a finite set B_m of vertices which, with positive probability, meets all the infinite clusters. Assume furthermore that $k(p) > 0$, that is, that there is at least one infinite cluster. Then, on the event where B_m meets all the infinite clusters, $N_{B_m}(0) \geq k(p)$ and $N_{B_m}(1) = 1$. Since $N_{B_m}(0) = N_{B_m}(1)$ almost surely, this implies that $k(p) \leq 1$. Thus, the only three possibilities left are $k(p) = +\infty$, $k(p) = 0$ and $k(p) = 1$. \square

In order to prove that $k(p) = 1$ for $p > p_c(d)$, it would remain to rule out the possibility that $k(p) = +\infty$. This is more difficult and we will not do it here.

4.3 The cluster of the origin in the subcritical phase

In the subcritical phase, that is, when $p < p_c(d)$, the cluster of the origin is almost surely finite. One may ask about the distribution of its size. There are at least two ways to measure the size of the cluster. One can consider the number of vertices that it contains, or the distance between the origin and the farthest vertex to which it is connected. These two quantities are of course related. It turns out to be convenient to consider the norm $\|\cdot\|_\infty$ on \mathbb{Z}^d , which is given by $\|(x_1, \dots, x_d)\|_\infty = \max(|x_1|, \dots, |x_d|)$.

Exercise 4.8 1. Check that the norm $\|\cdot\|_\infty$ is related to the usual Euclidean norm $\|\cdot\|$ by the inequalities

$$\forall x \in \mathbb{Z}^d, \quad \|x\|_\infty \leq \|x\| \leq \sqrt{d}\|x\|_\infty.$$

2. Check that, on \mathbb{L}^d , the inequality $|C(0)| \geq (2n)^d$ implies that there exists a vertex x such that $0 \leftrightarrow x$ and $\|x\|_\infty \geq n$.
3. Let us introduce the notation $S(n) = \{x \in \mathbb{Z}^d : \|x\|_\infty = n\}$. Prove that

$$\mathbb{P}_p(|C(0)| \geq n) \leq \mathbb{P}_p\left(0 \leftrightarrow S\left((n/2)^{\frac{1}{d}}\right)\right).$$

We define the *mean size of the cluster at the origin* as

$$\chi(p) = \mathbb{E}[|C(0)|].$$

There are many results about χ and the distribution of $|C(0)|$. We state one of them, which gives the correct order of magnitude of the size of the cluster at the origin.

Theorem 4.9 *Assume that $0 < p < p_c(d)$. Then there exists $\lambda(p) > 0$ such that, for all $n \geq 1$,*

$$\mathbb{P}_p(|C(0)| \geq n) \leq e^{-n\lambda(p)}.$$

Thus, the tail of the distribution of the size of $C(0)$ in the sub-critical phase is exponentially small. We will however not prove this fact. Let us state two weaker properties.

Proposition 4.10 *Assume that $0 < p < p_c(d)$.*

1. *There exists $\sigma(p) > 0$ such that, for all n ,*

$$\mathbb{P}_p(|C(0)| \geq n) \leq e^{-n^{\frac{1}{d}}\sigma(p)}.$$

2. $\chi(p) < +\infty$.

Exercise 4.11 *Deduce the first assertion of Proposition 4.10 from Theorem 4.9, and then the second assertion of Proposition 4.10 from the first.*

As a beautiful application of the BK inequality, we will only prove the first assertion of Proposition 4.10, assuming that the second is true.

Proposition 4.12 *Assume that $\chi(p) < +\infty$. Then there exists $\sigma(p) > 0$ such that, for all n ,*

$$\mathbb{P}_p(|C(0)| \geq n) \leq e^{-n^{\frac{1}{d}}\sigma(p)}.$$

Proof. Recall the notation of Exercise 4.8. According to the last assertion of this exercise, it suffices to prove that there exists a constant $\alpha(p) > 0$ such that

$$\mathbb{P}_p(0 \leftrightarrow S(n)) \leq e^{-n\alpha(p)}.$$

For each $x \in \mathbb{Z}^d$, let us denote by $\tau_p(0, x) = \mathbb{P}_p(0 \leftrightarrow x)$ the probability that 0 is connected to x . Let us also define, for all $n \geq 0$, the number N_n of vertices of $S(n)$ to which 0 is connected:

$$N_n = \#\{x \in S(n) : 0 \leftrightarrow x\}.$$

We have

$$\begin{aligned} \sum_{n \geq 0} \mathbb{E}_p[N_n] &= \sum_{n \geq 0} \sum_{x \in S(n)} \mathbb{P}_p(0 \leftrightarrow x) \\ &= \sum_{x \in \mathbb{Z}^d} \mathbb{P}_p(x \in C(0)) \\ &= \mathbb{E}_p[|C(0)|] \\ &= \chi(p). \end{aligned}$$

Since we are making the assumption that $\chi(p)$ is finite, we conclude that the series $\sum_{n \geq 0} \mathbb{E}_p[N_n]$ converges and in particular that $\mathbb{E}_p[N_n]$ tends to 0 as n tends to infinity.

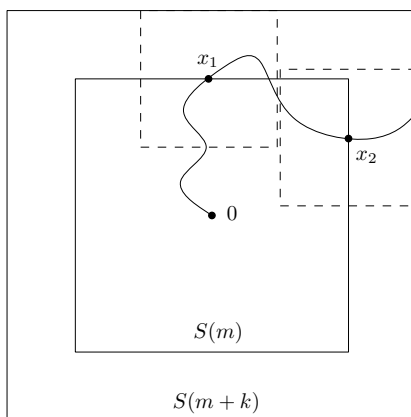


Figure 5: Different choices of x in $S(m)$ for whose the events $0 \leftrightarrow x$ and $x \leftrightarrow S(x, k)$ occur disjointly.

Let us now introduce another notation: for all $x \in \mathbb{Z}^d$ and all $k \geq 0$, we set $S(x, k) = \{y \in \mathbb{Z}^d : \|y - x\|_\infty = k\}$. In particular, $S(n)$ is simply $S(0, n)$. Now choose two integers $m, k \geq 0$. Assume that 0 is connected to $S(m+k)$. Then there exists a vertex x on $S(m)$ such that the events $0 \leftrightarrow x$ and $x \leftrightarrow S(x, k)$ occur disjointly. Hence, thanks to the BK

inequality,

$$\begin{aligned}
\mathbb{P}_p(0 \leftrightarrow S(m+k)) &\leq \mathbb{P}_p\left(\bigcup_{x \in S(m)} \{0 \leftrightarrow x\} \circ \{x \leftrightarrow S(x,k)\}\right) \\
&\leq \sum_{x \in S(m)} \mathbb{P}_p(\{0 \leftrightarrow x\} \circ \{x \leftrightarrow S(x,k)\}) \\
&\leq \sum_{x \in S(m)} \mathbb{P}_p(0 \leftrightarrow x) \mathbb{P}_p(x \leftrightarrow S(x,k)) \\
&= \sum_{x \in S(m)} \tau_p(0, x) \mathbb{P}_p(0 \leftrightarrow S(k)) \\
&= \mathbb{E}_p[N_m] \mathbb{P}_p(0 \leftrightarrow S(k)).
\end{aligned}$$

We have used the invariance by translation of the percolation process to say that

$$\mathbb{P}_p(x \leftrightarrow S(x, k)) = \mathbb{P}_p(0 \leftrightarrow S(k)).$$

Since $\mathbb{E}_p[N_n]$ tends to 0 as n tends to infinity, we can choose m large enough so that $\mathbb{E}_p[N_m] = \eta < 1$. Then, by using a simple arithmetic argument, one checks that

$$\mathbb{P}_p(0 \leftrightarrow S(n)) \leq \eta^{\frac{n}{m}-1},$$

which provides us with a bound of the expected form. □

5 The critical probability in two dimensions

In this last section on percolation, we use the results that we have collected to partially prove the following beautiful theorem.

Theorem 5.1 *For the percolation process on \mathbb{L}^2 , the critical probability is*

$$p_c(2) = \frac{1}{2}.$$

The proof is divided in two parts. We first prove that $p_c(2) \geq \frac{1}{2}$ and then that $p_c(2) \leq \frac{1}{2}$. For the first inequality, we use the following lemma.

Lemma 5.2 *Let A_1, \dots, A_m be increasing events with equal probability. Then*

$$\mathbb{P}_p(A_1) \geq 1 - (1 - \mathbb{P}_p(A_1 \cup \dots \cup A_m))^{\frac{1}{m}}.$$

Proof. We have $1 - \mathbb{P}_p(A_1 \cup \dots \cup A_m) = \mathbb{P}_p(A_1^c \cap \dots \cap A_m^c)$. Now the FKG inequality asserts that decreasing events are positively correlated. Hence,

$$1 - \mathbb{P}_p(A_1 \cup \dots \cup A_m) \geq \mathbb{P}_p(A_1^c) \dots \mathbb{P}_p(A_m^c) = \mathbb{P}_p(A_1^c)^m.$$

The results follows easily. \square

Proposition 5.3 *It is the case that $\theta_2(\frac{1}{2}) = 0$. In particular, $p_c(2) \geq \frac{1}{2}$.*

Proof. For each $n > 0$, let us denote by $A^l(n)$ (resp. $A^r(n)$, $A^t(n)$, $A^b(n)$) the event that some vertex of the left (resp. right, top, bottom) side of the box $T(n) = [0, n]^2$ is joined to infinity by an open path which crosses no other vertex of $T(n)$. The events $A^l(n)$, $A^r(n)$, $A^t(n)$, $A^b(n)$ are increasing and they have the same probability.

Let us assume that $\theta_2(\frac{1}{2}) > 0$. Then there exists almost surely an infinite cluster, so that

$$\mathbb{P}_{\frac{1}{2}}(A^l(n) \cup A^r(n) \cup A^t(n) \cup A^b(n)) \xrightarrow[n \rightarrow \infty]{} 1.$$

By Lemma 5.2, this implies that

$$\mathbb{P}_{\frac{1}{2}}(A^u(n)) \xrightarrow[n \rightarrow \infty]{} 1 \quad \text{for all } u = l, r, t, b.$$

Thus we may choose an N such that

$$\mathbb{P}_{\frac{1}{2}}(A^u(N)) > \frac{7}{8} \quad \text{for all } u = l, r, t, b.$$

Let us consider the dual box $T(n)_* = \{x + (\frac{1}{2}, \frac{1}{2}) : x \in T(n)\}$. Let $A_*^l(n)$ (resp. $A_*^r(n)$, $A_*^t(n)$, $A_*^b(n)$) be the event that some vertex of the left (resp. right, top, bottom) side of the dual box $T(n)_*$ is joined to infinity by a closed dual path which crosses no other vertex of $T(n)_*$. Since each edge of \mathbb{L}_*^2 is open or closed with probability $\frac{1}{2}$, we have

$$\mathbb{P}_{\frac{1}{2}}(A_*^u(N)) = \mathbb{P}_{\frac{1}{2}}(A^u(N)) > \frac{7}{8} \quad \text{for all } u = l, r, t, b.$$

Consider the event $A = A^l(n) \cap A^r(n) \cap A_*^t(n) \cap A_*^b(n)$. This event is depicted on Figure 6. The probability that A does not occur satisfies

$$\mathbb{P}_{\frac{1}{2}}(A^c) \leq \mathbb{P}_{\frac{1}{2}}(A^l(n)^c) + \mathbb{P}_{\frac{1}{2}}(A^r(n)^c) + \mathbb{P}_{\frac{1}{2}}(A_*^t(n)^c) + \mathbb{P}_{\frac{1}{2}}(A_*^b(n)^c) < \frac{1}{2},$$

so that $\mathbb{P}_p(A) > \frac{1}{2}$.

Now if A occurs, then the restriction to $\mathbb{L}^2 \setminus T(n)$ of the configuration contains at least two infinite open clusters, which meet $T(n)$ on its left and right side, and the restriction to $\mathbb{L}_*^2 \setminus T(n)_*$ of the dual configuration contains two infinite closed clusters, which meet $T(n)_*$ on its top and bottom side. It is impossible, for obvious topological reasons, that

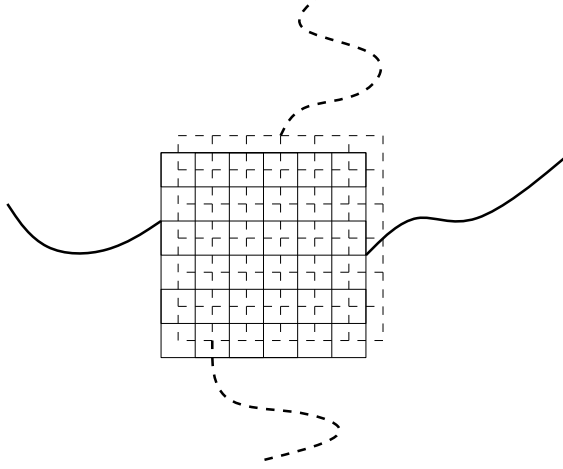


Figure 6: In this configuration, there must be at least two infinite clusters either in \mathbb{L}^2 or in \mathbb{L}_*^2 .

the two open clusters are joined inside $T(n)$ and at the same time the two dual closed clusters are joined inside $T(n)_*$.

Thus, when A occurs, the configuration or the dual configuration has more than one infinite cluster. By Theorem 4.6, this implies $\mathbb{P}_p(A) = 0$, in contradiction with our previous conclusions. Thus our assumption that $\theta_2(\frac{1}{2}) > 0$ was false. \square

We turn now to the proof of the other inequality.

Proposition 5.4 *The critical probability satisfies $p_c(2) \leq \frac{1}{2}$.*

Proof. Assume that $p < p_c(2)$. We will prove that there exists an infinite closed dual cluster with positive probability. This will imply $1 - p \geq p_c(2)$. Thus, $p < p_c(2) \Rightarrow 1 - p \geq p_c(2)$ and this would not be true if we had $p_c(2) > \frac{1}{2}$.

Let M be a positive integer. Let A_M denote the event that there exists an open path in \mathbb{L}^2 joining a vertex of the form $(k, 0)$ with $k < 0$ to a vertex of the form $(l, 0)$ with $l \geq M$, such that all the vertices of this path other than its endpoints lie strictly above the horizontal axis (see Figure 7).

If $(l, 0)$ is joined to $(k, 0)$ with $k < 0$, then the size of the cluster at $(l, 0)$ is at least l . Hence,

$$\begin{aligned} \mathbb{P}_p(A_M) &\leq \mathbb{P}_p\left(\bigcup_{l \geq M} \{\exists k < 0 : (l, 0) \leftrightarrow (k, 0)\}\right) \\ &\leq \sum_{l=M}^{\infty} \mathbb{P}_p(|C((l, 0))| \geq l) \\ &= \sum_{l=M}^{\infty} \mathbb{P}_p(|C(0)| \geq l), \end{aligned}$$

where we have used the fact that the distribution of the size of the cluster is the same at any vertex.

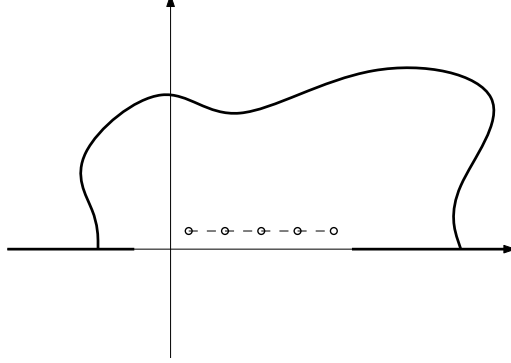


Figure 7: The event A_M .

By Proposition 4.10, $\chi(p) < +\infty$ because we are assuming $p < p_c(d)$. Hence, the last series is convergent and, if we choose M large enough, we may assume that $\mathbb{P}_p(A_M) \leq \frac{1}{2}$.

We claim that if A_M does not occur, then there exists an infinite closed cluster in the dual lattice. Indeed, let L be the dual segment consisting in the vertices $\{(m + \frac{1}{2}, \frac{1}{2}) : 0 \leq m < M\}$. Let $C(L)_*$ be the set of the dual vertices which are joined by a closed path to a vertex of L . If $|C(L)_*| < +\infty$, then, by Proposition 2.24, $C(L)_*$ is surrounded by an open closed circuit in \mathbb{L}^2 , and the existence of such a closed circuit guarantees that A_M occurs. Hence,

$$\mathbb{P}_p(|C(L)_*| < +\infty) \leq \mathbb{P}_p(A_M) \leq \frac{1}{2}.$$

This implies that $\mathbb{P}_p(|C(L)_*| = +\infty) > 0$. Hence, the dual percolation process, whose edges are closed with probability $1 - p$, admits an infinite closed cluster with positive probability. This implies that $1 - p \geq p_c(d)$, as expected. \square

Part II

The Ising model

The Ising model was invented by Lenz in order to provide a mathematical model for the behaviour of ferromagnetic metals in presence of a magnetic field. It has been known for a long time that a piece of iron which has been exposed temporarily to a magnetic field keeps a memory of this exposition and produces itself a magnetic field, provided it is not too hot. There is a critical temperature, called the Curie temperature, above which the piece of iron loses its magnetization.

The Ising model interprets this phenomenon by saying that the piece of iron contains many tiny magnets which an exterior field can tend to align in a given direction. More precisely, the exterior magnetic field acts on the magnets which are close to the surface of the piece of iron, by aligning them, and then the magnets inside the piece have interest, from an energetic point of view, to be aligned to the extent possible with their neighbours. Of course, the thermal agitation perturbs this tendency to organization and gives each magnet a somewhat random behaviour. The point is that above a certain critical temperature, the thermal effect overrides the magnetic organization.

There are several differences between the Ising model and percolation. Firstly, the Ising configurations live on the vertices of a graph rather than on the edges. Then, the values of a configuration at different vertices are not independent. Also, it is much more difficult to define the Ising model on an infinite graph than on a finite graph. We will discuss infinite volume limits, for this is the only way to see phase transitions occur, but we will not really define the Ising model on an infinite graph, not even on \mathbb{L}^d .

6 The probability measure

Let $G = (V, E)$ be a finite graph. If x, y are vertices, recall that we write $x \sim y$ if the pair $\{x, y\}$ is an edge. The configuration space of the Ising model on G is the set $\Sigma = \{-1, 1\}^V$ of all functions from V to $\{-1, 1\}$. The value of the configuration at a vertex is usually called a *spin*. This corresponds to the idea of a small magnet at each vertex, which can point either up or down. In contrast to the percolation model, the value of a configuration at each site is not independent of its values at other sites. Instead, to each configuration is associated a certain energy, which determines the probability that this particular configuration occurs. The function which to each configuration associates its energy is called the Hamiltonian.

Definition 6.1 *The Hamiltonian of the Ising model on $G = (V, E)$ is the function $H : \Sigma \rightarrow \mathbb{R}$ defined by*

$$\forall \sigma \in \Sigma, H(\sigma) = - \sum_{x \sim y} \sigma(x)\sigma(y).$$

For the sake of precision, let us emphasize that each edge $\{x, y\}$ appears once (and not twice) in the sum which defines the Hamiltonian. In particular, the lowest possible value for H is $-|E|$ and it is achieved only by the two constant configurations which associate the same spin to each vertex. The highest possible value for H depends on the geometry of the graph G .

Exercise 6.2 Find a geometric condition on a graph G which is necessary and sufficient for the highest possible value of the Ising Hamiltonian to be $|E|$.

Definition 6.3 Let β be a non-negative real number. We define the probability measure \mathbb{Q}_β on Σ by setting, for all $\sigma \in \Sigma$,

$$\mathbb{Q}_\beta(\sigma) = \frac{e^{-\beta H(\sigma)}}{\sum_{\sigma \in \Sigma} e^{-\beta H(\sigma)}}.$$

The denominator in this expression is usually denoted by

$$Z(\beta) = \sum_{\sigma \in \Sigma} e^{-\beta H(\sigma)}$$

and it is called the partition function.

The real β is usually called the *inverse temperature*. It plays the role of the coefficient $\frac{1}{kT}$ which appears in physics, where k is the Boltzmann constant and T the temperature. For $\beta = 0$, which corresponds to an infinite temperature, the measure \mathbb{Q}_β is simply the uniform measure on Σ : the Hamiltonian has no effect at all. This corresponds to the physical situation where the magnetic interaction of spins is negligible with respect to thermal agitation. At the other extreme, when β is very large, configurations with a high energy, that is, on which H takes a large value, are strongly penalized. There is a very strong tendency to align spins, and thermal agitation has only a minor effect.

Exercise 6.4 The Ising model has an important symmetry property : each configuration has the same energy as the opposite configuration. Check this, and use it to prove that for all vertex x , $\mathbb{Q}_\beta(\sigma(x) = 1) = \mathbb{Q}_\beta(\sigma(x) = -1) = \frac{1}{2}$.

7 First examples

Let us consider the Ising model on some very simple graphs.

7.1 The segment

Choose an integer $n \geq 2$. Consider the graph whose vertices are the integers $1, \dots, n$ and whose edges are $\{1, 2\}, \{2, 3\}, \dots, \{n-1, n\}$. The Hamiltonian in this case is

$$H(\sigma) = - \sum_{k=1}^{n-1} \sigma(k)\sigma(k+1).$$

A configuration σ is completely determined by $\sigma(1)$ and the $n-1$ signs $\varepsilon(k) = \sigma(k)\sigma(k+1)$ for $k \in \{1, \dots, n-1\}$. The following proposition explicits the model in terms of these variables.

Proposition 7.1 *Under \mathbb{Q}_β , the random variables $\sigma(1), \varepsilon(1), \varepsilon(2), \dots, \varepsilon(n-1)$ are independent. Moreover, $\varepsilon(1), \varepsilon(2), \dots, \varepsilon(n-1)$ are identically distributed with distribution*

$$\mathbb{Q}_\beta(\varepsilon_1 = 1) = \frac{e^\beta}{e^\beta + e^{-\beta}}, \quad \mathbb{Q}_\beta(\varepsilon_1 = -1) = \frac{e^{-\beta}}{e^\beta + e^{-\beta}}.$$

Exercise 7.2 *Prove Proposition 7.1. Prove that $Z(\beta) = 2^n (\cosh \beta)^{n-1}$.*

Thus, in this case, the Ising model can be expressed easily as a function of a collection of independent random variables.

7.2 The complete graph

Choose again an integer $n \geq 2$. Consider the graph whose vertices are the integers $1, \dots, n$ and whose edges are all the pairs $\{k, l\}$ with $k \neq l$ between 1 and n . Thus, all possible edges are present. The Hamiltonian has in this case a very nice expression. If we call $S(\sigma) = \sum_{k=1}^n \sigma(k)$ the sum of the spins, we have

$$H(\sigma) = - \sum_{1 \leq k < l \leq n} \sigma(k)\sigma(l) = -\frac{1}{2} \left(S(\sigma)^2 - \sum_{k=1}^n \sigma(k)^2 \right) = -\frac{1}{2} (S(\sigma)^2 - n).$$

The possible values of $|S(\sigma)|$ are $0, 1, \dots, n$. For each $s \in \{0, \dots, n\}$, let us define $\Sigma_s = \{\sigma \in \Sigma : |S(\sigma)| = s\}$. We have a partition $\Sigma = \Sigma_0 \sqcup \dots \sqcup \Sigma_n$. For each s , let ν_s denote the uniform probability measure on Σ_s .

Proposition 7.3 *The measure \mathbb{Q}_β can be written as*

$$\mathbb{Q}_\beta = \frac{1}{Z_0(\beta)} \sum_{s=0}^n e^{\frac{\beta}{2}s^2} \nu_s,$$

with $Z_0(\beta) = \sum_{s=0}^n e^{\frac{\beta}{2}s^2}$.

Exercise 7.4 *Check that $S(\sigma)$ has the same parity as n , so that half of the spaces Σ_s , $s \in \{0, \dots, n\}$ are empty.*

8 The phase transition

For the Ising model on a finite graph, every event has a probability which depends smoothly on β . There is no phase transition in a finite world. In order to observe phase transitions, we must consider infinite graphs. Since it is difficult to define the Ising model on infinite graphs, we will instead consider sequences of graphs which become larger and larger. The typical quantity that we are going to consider is the following: consider a typical configuration in a very large box, conditioned to have all spins up on the boundary of the box. Does the conditioning have an influence on the behaviour of the spin at the center of the box? Is it more likely to point up than without this boundary condition? With a finite box, even very large, the answer is always yes, the probability that the spin at the center of the box points up is strictly larger than $\frac{1}{2}$. However, as the size of the box increases, this probability decreases. Its limit may either be $\frac{1}{2}$, in which case we say that the spin at the center of the (infinite) box is not influenced by the boundary, or it may be strictly larger than $\frac{1}{2}$. This depends of course on the parameter β and we will see that there is a critical value above which one observes one behaviour and below which one observes the other.

8.1 Magnetization

For each integer $n \geq 1$, let us denote by Λ_n^d , or simply Λ_n , the subgraph of \mathbb{L}^d with vertices $\{-n, \dots, n\}^d$ and all the edges of \mathbb{L}^d which join two such vertices. Let us also denote by $\partial\Lambda_n$ the set of vertices located on the boundary of Λ_n , that is, the set of vertices of Λ_n which are joined in \mathbb{L}^d to at least one vertex of $\mathbb{Z}^d \setminus \{-n, \dots, n\}^d$.

Let us denote by Σ_n the configuration space of the Ising model on Λ_n , and by $\mathbb{Q}_{\beta,n}$ the Ising probability measure. We define now the set Σ_n^+ of configurations with boundary conditions 1 as

$$\Sigma_n^+ = \{\sigma \in \Sigma_n : \forall x \in \partial\Lambda_n, \sigma(x) = 1\}.$$

We define $\mathbb{Q}_{\beta,n}^+$ as the conditioned measure on Σ_n^+ , that is, for all $\sigma \in \Sigma_n$,

$$\mathbb{Q}_{\beta,n}^+(\{\sigma\}) = \frac{\mathbb{1}_{\sigma \in \Sigma_n^+} \mathbb{Q}_{\beta,n}(\{\sigma\})}{\mathbb{Q}_{\beta,n}(\Sigma_n^+)}.$$

The measure $\mathbb{Q}_{\beta,n}$ modelizes the distribution of the spins inside a piece of metal when all spins located on the surface of the piece are aligned by some exterior magnetic field.

For the Ising model without boundary conditions, we have $\mathbb{Q}_{\beta,n}(\sigma(0) = 1) = \frac{1}{2}$, because of the symmetry of the Ising model (see Exercise 6.4). Now, in the presence of the boundary conditions, more spins tend to take the value 1, at least close to the boundary of the box. It may or may not be the case that this effect is felt through the entire box, but in any case it is likely that

$$\mathbb{Q}_{\beta,n}^+(\sigma(0) = 1) \geq \frac{1}{2}. \quad (4)$$

We will prove later that this inequality holds and is in fact a strict inequality.

Now, as n grows to infinity, the effect of boundary conditions becomes weaker and weaker at the origin. It is thus also likely that $\mathbb{Q}_{\beta,n}^+(\sigma(0) = 1)$ is decreasing with respect to n .

Definition 8.1 *The magnetization of the Ising model on \mathbb{L}^d is the real number $M(\beta)$ defined by*

$$M(\beta) = \liminf_{n \rightarrow \infty} \mathbb{Q}_{\beta,n}^+(\sigma(0) = 1) - \frac{1}{2}.$$

The \liminf is only here to avoid any difficulty with the existence of the limit, but we shall prove that it is indeed a limit.

A magnetization equal to 0 means that the interaction between neighbouring spins is too weak for the origin to feel, in the limit where the size of the box tends to infinity, the effect of the boundary conditions. For $\beta = 0$, it is clear that $M(0) = 0$. On the other hand, a positive magnetization means that the interaction between spins is strong enough to communicate the effect of the boundary conditions to a point which is macroscopically far from the boundary of the box. In the $\beta \rightarrow \infty$ limit, the measure $\mathbb{Q}_{n,\beta}^+$ charges almost exclusively the configuration with all spins aligned, equal to +1, and it seems plausible that $\lim_{\beta \rightarrow +\infty} M(\beta) = 1$. In fact, this is true only if $d \geq 2$, and to prove this is by no means an easy task.

The main question that we are going to discuss is the following : for which values of β is it the case that $M(\beta) > 0$? The answer depends on the dimension d .

Theorem 8.2 (Phase transition for the Ising model) *If $d = 1$, then $M(\beta) = 0$ for all $\beta \in [0, +\infty)$. If $d \geq 2$, there exists $\beta_c > 0$ such that $M(\beta) = 0$ for $0 \leq \beta < \beta_c$ and $M(\beta) > 0$ for $\beta > \beta_c$.*

In particular, for $d \geq 2$, the magnetization is not an analytic function of β . Hence, we may speak of a phase transition.

8.2 The one-dimensional case

In the case of the line, it is possible to compute everything directly.

Proposition 8.3 *Assume that $d = 1$. Choose $\beta \in [0, +\infty)$. Set $p = \frac{e^\beta}{e^\beta + e^{-\beta}}$, so that $p \in [\frac{1}{2}, 1)$. Then, for all $n \geq 1$,*

$$\mathbb{Q}_{\beta,n}^+(\sigma(0) = 1) - \frac{1}{2} = \frac{(2p - 1)^n}{1 + (2p - 1)^{2n}}.$$

In particular, $M(\beta) = 0$.

The main technical tool is the following.

Lemma 8.4 For each $k \in \{1, \dots, n\}$, set $\varepsilon_k = \sigma(k-1)\sigma(k)$ and $\varepsilon_{-k} = \sigma(-k+1)\sigma(-k)$. Then, under $\mathbb{Q}_{\beta,n}^+$, the random variables $\varepsilon_1, \dots, \varepsilon_n, \varepsilon_{-1}, \dots, \varepsilon_{-n}$ have the distribution of $2n$ independent random variables with distribution $p\delta_1 + (1-p)\delta_{-1}$ conditioned by the event that their product is equal to 1.

Proof. We have $H(\sigma) = -\sum_{k=1}^n (\varepsilon_k + \varepsilon_{-k})$. Let us choose $\alpha_1, \alpha_{-1}, \dots, \alpha_n, \alpha_{-n} \in \{-1, 1\}$. Since $\sigma(n) = \sigma(-n)\varepsilon_{-n} \dots \varepsilon_{-1}\varepsilon_1 \dots \varepsilon_n$ and since $\sigma(n) = \sigma(-n) = 1$ almost surely under $\mathbb{Q}_{\beta,n}^+$, the probability that $\varepsilon_k = \alpha_k$ for all k is zero unless the product of the α_k 's is equal to 1. In this case,

$$\mathbb{Q}_{\beta,n}^+(\varepsilon_1 = \alpha_1, \varepsilon_{-1} = \alpha_{-1}, \dots, \varepsilon_n = \alpha_n, \varepsilon_{-n} = \alpha_{-n}) \propto \prod_{k=1}^n e^{\beta\alpha_k} e^{\beta\alpha_{-k}},$$

and the result follows. \square

Proof of Proposition 8.3 – Let $(\gamma_k)_{k \in \mathbb{Z}}$ be i.i.d. random variables with distribution $p\delta_1 + (1-p)\delta_{-1}$. Since $\sigma(0) = \varepsilon_1 \dots \varepsilon_n \sigma(n) = \varepsilon_1 \dots \varepsilon_n$, and thanks to the lemma above,

$$\begin{aligned} \mathbb{Q}_{\beta,n}^+(\sigma(0) = 1) &= \mathbb{P}(\gamma_1 \dots \gamma_n = 1 | \gamma_1 \dots \gamma_n \gamma_{-1} \dots \gamma_{-n} = 1) \\ &= \frac{\mathbb{P}(\gamma_1 \dots \gamma_n = 1, \gamma_{-1} \dots \gamma_{-n} = 1)}{\mathbb{P}(\gamma_1 \dots \gamma_n \gamma_{-1} \dots \gamma_{-n} = 1)} \\ &= \frac{q_n^2}{q_{2n}}, \end{aligned}$$

where we have set, for all $m \geq 1$, $q_m = \mathbb{P}(\gamma_1 \dots \gamma_m = 1)$. Now a direct computation shows that

$$q_m = \sum_{k=0}^{\lfloor \frac{m}{2} \rfloor} \binom{m}{2k} p^{m-2k} (1-p)^k = \frac{1}{2} (1 + (2p-1)^m).$$

The result follows easily. \square

Let us now compute the partition function $Z_n^+(\beta) = \sum_{\sigma \in \Sigma_n^+} e^{-\beta H(\sigma)}$.

Proposition 8.5 The partition function is given by

$$Z_n^+(\beta) = (2 \cosh \beta)^{2n} \frac{1 + (2p-1)^{2n}}{2}.$$

In particular, the free energy is given by

$$F^+(\beta) = \lim_{n \rightarrow \infty} \frac{1}{2n} \log Z_n^+(\beta) = 2 \cosh \beta.$$

Proof. We use again the variables $\varepsilon_1, \varepsilon_{-1}, \dots, \varepsilon_n, \varepsilon_{-n}$. We have

$$Z_n^+(\beta) = \sum_{\substack{\varepsilon_1, \varepsilon_{-1}, \dots, \varepsilon_n, \varepsilon_{-n} \in \{-1, +1\} \\ \varepsilon_1 \dots \varepsilon_n \varepsilon_{-1} \varepsilon_{-n} = 1}} e^{\beta \varepsilon_1} \dots e^{\beta \varepsilon_n} e^{\beta \varepsilon_{-1}} \dots e^{\beta \varepsilon_{-n}} = z_{2n}^+,$$

where we set, for all $m \geq 0$,

$$z_m^\pm = \sum_{\substack{\gamma_1, \dots, \gamma_m \in \{-1, +1\} \\ \gamma_1 \dots \gamma_m = \pm 1}} e^{\beta \gamma_1} \dots e^{\beta \gamma_m}.$$

We make the convention $z_0^+ = 1$ and $z_0^- = 0$. The sequences $(z_m^+)_{m \geq 0}$ and $(z_m^-)_{m \geq 0}$ satisfy the recurrence system

$$z_{m+1}^+ = e^\beta z_m^+ + e^{-\beta} z_m^-, \quad z_{m+1}^- = e^{-\beta} z_m^+ + e^\beta z_m^-,$$

which can be rewritten using matrices as

$$\begin{pmatrix} z_m^+ \\ z_m^- \end{pmatrix} = \begin{pmatrix} e^\beta & e^{-\beta} \\ e^{-\beta} & e^\beta \end{pmatrix}^m \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

The 2×2 matrix which appears here is equal to $2 \cosh \beta \begin{pmatrix} p & 1-p \\ 1-p & p \end{pmatrix}$. The eigenvalues of the last matrix are 1 and $2p-1$, with respective eigenvectors $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. It is now easy to obtain a closed formula for z_m^+ . \square

9 The random cluster model

In order to analyze the Ising model in two dimensions or more, we introduce a new model, the random cluster model. It is an edge model, similar to percolation, indeed a generalization of percolation. It is sometimes called FK-percolation, in reference to Fortuin and Kasteleyn, who first introduced it. The interest of this model for the study of the Ising model comes from two facts. The first is that the random cluster model can be coupled to the Ising model, in such a nice way that the magnetization of the Ising model can be expressed in terms of the percolation probability for the random cluster model. The second is that there are relations of stochastic domination between percolation and the random cluster model, for appropriate values of the parameters. Thus, our knowledge of percolation will allow us to prove the existence of a phase transition for the random cluster model, hence for the Ising model.

9.1 The random cluster measure

Let $G = (V, E)$ be a finite graph. The random cluster model is a probability measure on $\Omega = \{0, 1\}^E$ which depends on two real parameters $p \in [0, 1]$ and $q > 0$. We denote it

by $\mathbb{R}_{p,q}$. For each configuration $\omega \in \Omega$, we denote by $\text{cl}(\omega)$ the number of clusters of ω . Recall that $K(\omega)$ is the set of edges which are open in ω . The measure $\mathbb{R}_{p,q}$ is defined as follows:

$$\forall \omega \in \Omega, \quad \mathbb{R}_{p,q}(\omega) = \frac{1}{Z(p,q)} p^{|K(\omega)|} (1-p)^{|E \setminus K(\omega)|} q^{\text{cl}(\omega)}, \quad (5)$$

where the partition function $Z(p,q)$ is the unique positive constant which makes $\mathbb{R}_{p,q}$ a probability measure.

For $q = 1$, $\mathbb{R}_{p,1} = \mathbb{P}_p$ is the percolation measure with parameter p . For $q < 1$, configurations with many clusters, hence few edges, are penalized. We will focus on the case where $q \geq 1$, indeed $q = 2$. In this case, the configurations with few clusters are penalized. Thus, a typical configuration under $\mathbb{R}_{p,2}$ should have fewer edges as a typical configuration under \mathbb{P}_p . We will ground these intuitions on rigorous results. To start with, we prove that, for $q \geq 1$, the measure $\mathbb{R}_{p,q}$ satisfies the FKG inequality.

Proposition 9.1 *For all $p \in [0, 1]$ and all $q \in [1, +\infty)$, the measure $\mathbb{R}_{p,q}$ satisfies the FKG condition (see Proposition 3.7).*

Proof. Let ω_1 and ω_2 be two configurations. We need to prove that

$$\mathbb{R}_{p,q}(\omega_1 \vee \omega_2) \mathbb{R}_{p,q}(\omega_1 \wedge \omega_2) \geq \mathbb{R}_{p,q}(\omega_1) \mathbb{R}_{p,q}(\omega_2).$$

Let K_1 (resp. K_2) denote the set of edges which are open in ω_1 (resp. ω_2). According to (5), we need to prove that

$$p^{|K_1 \cup K_2| + |K_1 \cap K_2|} (1-p)^{2|E| - |K_1 \cup K_2| - |K_1 \cap K_2|} q^{\text{cl}(\omega_1 \vee \omega_2) + \text{cl}(\omega_1 \wedge \omega_2)} \geq p^{|K_1| + |K_2|} (1-p)^{2|E| - |K_1| - |K_2|} q^{\text{cl}(\omega_1) + \text{cl}(\omega_2)}.$$

By the result of Exercise 3.9, p and $(1-p)$ appear with the same power on each side of the inequality to prove. Thus, under the assumption $q \geq 1$, the problem reduces to proving that the following purely geometric inequality holds:

$$\text{cl}(\omega_1 \vee \omega_2) + \text{cl}(\omega_1 \wedge \omega_2) \geq \text{cl}(\omega_1) + \text{cl}(\omega_2).$$

For this, let us construct an abstract graph $H = (S, A)$ (which is different from the graph underlying our random cluster model) as follows. The set S of vertices of this graph is the disjoint union of two sets S_1 and S_2 , where S_1 is the set of clusters of ω_1 and S_2 the set of clusters of ω_2 . Now we join two elements of S by an edge if and only if, as clusters of G , they have at least one common vertex.

Observe that two vertices of S_1 (or two vertices of S_2) cannot be joined by an edge, for by definition two distinct clusters of ω_1 (or ω_2) do not intersect. Hence, each edge of H joins a vertex of S_1 to a vertex of S_2 .

By definition, $\text{cl}(\omega_1) + \text{cl}(\omega_2) = |S|$, the number of vertices of the graph H .

Let us count $\text{cl}(\omega_1 \vee \omega_2)$ in terms of H . Observe that each vertex of G determines two vertices of H . Indeed, let x be a vertex of G . Then x belongs to a cluster of ω_1 , say $s_1(x)$,

and also to a cluster $s_2(x)$ of ω_2 . Moreover, $s_1(x)$ and $s_2(x)$ share at least the vertex x , so that, as vertices of H , they are joined by an edge. Finally, each vertex of G determines an edge of H . Two vertices x and y of G determine two edges of H which have a common endpoint if and only if x and y are in the same cluster of ω_1 or in the same cluster of ω_2 .

Now let x and y be two vertices of G . They are in the same cluster of $\omega_1 \vee \omega_2$ if and only if there exists a chain of vertices $x = x_1, x_2, \dots, x_n = y$ such that for all $i \in \{1, \dots, n-1\}$, the vertices x_i and x_{i+1} are either in the same cluster of ω_1 or in the same cluster of ω_2 . In abstract terms, the relation “to be in the same cluster for $\omega_1 \vee \omega_2$ ” is the transitive closure of the relation “to be in the same cluster for ω_1 or in the same cluster for ω_2 ”. In terms of the graph H , this means that x and y determine two edges which are joined in H by a chain of edges, each sharing an endpoint with the next. Thus, x and y are in the same cluster for $\omega_1 \vee \omega_2$ if and only if the edges associated with x and y belong to the same connected component of H . Finally, $\text{cl}(\omega_1 \vee \omega_2)$ is the number of connected components of the graph H .

Let us finally count $\text{cl}(\omega_1 \wedge \omega_2)$ in terms of H . Let x and y be two vertices of G . Assume that they are in the same cluster of $\omega_1 \wedge \omega_2$. Then they are connected by ω_1 and ω_2 (beware that the converse is false). Thus, $s_1(x) = s_1(y)$ and $s_2(x) = s_2(y)$, so that x and y determine the same edge of H . Hence, $\text{cl}(\omega_1 \wedge \omega_2) \geq |A|$, the number of edges of H .

Now it is a general fact, for any finite graph, that

$$\# \text{ edges} + \# \text{ connected components} \geq \# \text{ vertices}.$$

This can be proved by induction on the number of edges, and completes the argument. \square

9.2 Inequalities between random cluster model and percolation

The fact that the random cluster model satisfies the FKG inequality allows us to prove relations of stochastic domination between the random cluster model and the percolation.

Proposition 9.2 *Choose $p_1, p_2 \in [0, 1]$ and $q_1, q_2 \in [1, +\infty)$.*

1. *If $p_1 \leq p_2$ and $q_1 \geq q_2$, then $\mathbb{R}_{p_1, q_1} \leq_{\text{st}} \mathbb{R}_{p_2, q_2}$.*
2. *If $\frac{p_1}{q_1(1-p_1)} \geq \frac{p_2}{q_2(1-p_2)}$ and $q_1 \geq q_2$, then $\mathbb{R}_{p_1, q_1} \geq_{\text{st}} \mathbb{R}_{p_2, q_2}$.*

Proof. Let X be an increasing random variable. We compute $\mathbb{R}_{p_2, q_2}(X)$.

$$\begin{aligned} \mathbb{R}_{p_2, q_2}(X) &= \frac{1}{Z(p_2, q_2)} \sum_{\omega \in \Omega} X(\omega) p_2^{|K(\omega)|} (1-p_2)^{|E \setminus K(\omega)|} q_2^{\text{cl}(\omega)} \\ &= \left(\frac{1-p_2}{1-p_1} \right)^{|E|} \frac{1}{Z(p_2, q_2)} \sum_{\omega \in \Omega} X(\omega) \underbrace{\left(\frac{\frac{p_2}{1-p_2}}{\frac{p_1}{1-p_1}} \right)^{|K(\omega)|} \left(\frac{q_2}{q_1} \right)^{\text{cl}(\omega)}}_{Y(\omega)} p_1^{|K(\omega)|} (1-p_1)^{|E \setminus K(\omega)|} q_1^{\text{cl}(\omega)} \\ &= \left(\frac{1-p_2}{1-p_1} \right)^{|E|} \frac{Z(p_1, q_1)}{Z(p_2, q_2)} \mathbb{R}_{p_1, q_1}(XY). \end{aligned}$$

Applying this formula with $X = 1$ gives us the value of $\mathbb{R}_{p_1, q_1}(Y)$ and we find

$$\mathbb{R}_{p_2, q_2}(X) = \frac{\mathbb{R}_{p_1, q_1}(XY)}{\mathbb{R}_{p_1, q_1}(Y)}.$$

Now $|K(\omega)|$ is an increasing function of ω and $\text{cl}(\omega)$ is a decreasing function of ω . Hence, under the first set of assumptions of p_1, p_2, q_1, q_2 , Y is an increasing random variable. Hence, by the FKG inequality, $\mathbb{R}_{p_2, q_2}(X) \geq \mathbb{R}_{p_1, q_1}(X)$. This proves the first assertion.

In order to see what happens under the second set of assumptions, let us rewrite $Y(\omega)$ as

$$Y(\omega) = \left(\frac{\frac{p_2}{q_2(1-p_2)}}{\frac{p_1}{q_1(1-p_1)}} \right)^{|K(\omega)|} \left(\frac{q_2}{q_1} \right)^{\text{cl}(\omega) + |K(\omega)|}.$$

The point of this way of writing Y is that the exponent $\text{cl}(\omega) + |K(\omega)|$ is now also an increasing function of ω . Thus, under the second set of assumptions, Y is a decreasing random variable. Hence, the FKG implies that $\mathbb{R}_{p_2, q_2}(X) \leq \mathbb{R}_{p_1, q_1}(X)$. This finishes the proof. \square

Exercise 9.3 1. Check that $\text{cl}(\omega) + |K(\omega)|$ is an increasing function of ω .
2. Check that Proposition 9.2 implies the following inequalities, for all $p \in [0, 1]$:

$$\mathbb{P}_{\frac{p}{2-p}} \leq_{\text{st}} \mathbb{R}_{p, 2} \leq_{\text{st}} \mathbb{P}_p.$$

9.3 Coupling the random cluster and Ising models

Let $G = (V, E)$ be a finite graph. We want to couple the Ising model and the random cluster model on G , that is, we want to find a probability measure on $\Sigma \times \Omega = \{-1, 1\}^V \times \{0, 1\}^E$ whose marginals are the Ising measure for some parameter β on Σ and the random cluster measure for some parameters p and q on Ω .

Definition 9.4 Choose $p \in [0, 1]$. We define a probability measure \mathbb{S}_p on $\Sigma \times \Omega$ by setting, for all (σ, ω) ,

$$\mathbb{S}_p(\sigma, \omega) = \frac{1}{Z} \prod_{e \in E, e = \{x, y\}} \left((1-p) \mathbb{1}_{\omega(e)=0} + p \mathbb{1}_{\omega(e)=1} \mathbb{1}_{\sigma(x)=\sigma(y)} \right),$$

where, as usual, Z is the unique real constant which makes \mathbb{S}_p a probability measure.

Proposition 9.5 The marginal distribution on Σ of the measure \mathbb{S}_p is the Ising measure \mathbb{Q}_β with $p = 1 - e^{-2\beta}$: for all $\sigma \in \Sigma$,

$$\sum_{\omega \in \Omega} \mathbb{S}_p(\sigma, \omega) = \mathbb{Q}_\beta(\sigma).$$

Proof. Choose $\sigma \in \Sigma$. We have

$$\begin{aligned}
\sum_{\omega \in \Omega} \mathbb{S}_p(\sigma, \omega) &= \frac{1}{Z} \prod_{e \in E, e=\{x,y\}} \sum_{\omega_e \in \{0,1\}} ((1-p)\mathbb{1}_{\omega(e)=0} + p\mathbb{1}_{\omega(e)=1}\mathbb{1}_{\sigma(x)=\sigma(y)}) \\
&= \frac{1}{Z} \prod_{e \in E, e=\{x,y\}} (\mathbb{1}_{\sigma(x)=\sigma(y)} + (1-p)\mathbb{1}_{\sigma(x) \neq \sigma(y)}) \\
&= \frac{1}{Z} \prod_{e \in E, e=\{x,y\}} \sqrt{1-p} e^{-\frac{1}{2} \log(1-p)\sigma(x)\sigma(y)} \\
&= \frac{1}{Z'} \exp\left(-\frac{\log(1-p)}{2} \sum_{\{x,y\} \in E} \sigma(x)\sigma(y)\right),
\end{aligned}$$

where $Z' = (1-p)^{-\frac{|E|}{2}} Z$. We recognize a measure which is proportional, hence equal to the Ising model on G with $\beta = -\frac{1}{2} \log(1-p)$, that is, $p = 1 - e^{-2\beta}$. \square

Let us determine the marginal of \mathbb{S}_p on Ω .

Proposition 9.6 *The marginal distribution on Ω of the measure \mathbb{S}_p is the random cluster measure $\mathbb{R}_{p,2}$: for all $\omega \in \Omega$,*

$$\sum_{\sigma \in \Sigma} \mathbb{S}_p(\sigma, \omega) = \mathbb{R}_{p,2}(\omega).$$

Proof. Choose $\omega \in \Omega$. We have

$$\sum_{\sigma \in \Sigma} \mathbb{S}_p(\sigma, \omega) = \frac{1}{Z} \sum_{\sigma \in \Sigma} \prod_{e \in E, e=\{x,y\}} ((1-p)\mathbb{1}_{\omega(e)=0} + p\mathbb{1}_{\omega(e)=1}\mathbb{1}_{\sigma(x)=\sigma(y)}).$$

Only those σ which assign the same spin to all vertices of each cluster of ω give a non-zero contribution to this sum. Moreover, each such σ gives the same contribution $p^{\#\{e \in E: \omega(e)=1\}} (1-p)^{\#\{e \in E: \omega(e)=0\}}$. Since each spin can take 2 distinct values, there are exactly $2^{\text{cl}(\omega)}$ configurations σ which contribute. This is the expected result. \square

Finally, let us describe the conditional distribution of the first component of (σ, ω) given the second under \mathbb{S}_p . The next statement is in fact a consequence of the argument that we have just used in the last proof.

Proposition 9.7 *Choose $\pi \in \Omega$. Then the conditional distribution of σ given $\omega = \pi$ under \mathbb{S}_p is the uniform distribution on all spin configurations which assign the same spin to all vertices of each cluster of π .*

The crucial result, which provides us with an effective link between the Ising model and the random cluster model, is the following.

Proposition 9.8 Choose β and p such that $p = 1 - e^{-2\beta}$. For all $x, y \in V$, it is the case that

$$\mathbb{Q}_\beta(\sigma(x) = \sigma(y)) - \frac{1}{2} = \frac{1}{2} \mathbb{R}_{p,2}(x \leftrightarrow y).$$

Proof. Since \mathbb{Q}_β is a marginal of \mathbb{S}_p , we have

$$\begin{aligned} \mathbb{Q}_\beta(\sigma(x) = \sigma(y)) - \frac{1}{2} &= \sum_{\sigma, \omega} \left(\mathbb{1}_{\sigma(x) = \sigma(y)} - \frac{1}{2} \right) \mathbb{S}_p(\sigma, \omega) \\ &= \sum_{\omega} \mathbb{R}_{p,2}(\omega) \sum_{\sigma} \mathbb{S}_p(\sigma | \omega) \left(\mathbb{1}_{\sigma(x) = \sigma(y)} - \frac{1}{2} \right), \end{aligned}$$

where $\mathbb{S}_p(\sigma | \omega)$ is a shorthand for the conditional probability that the first component is σ given that the second is ω . Now, for each ω , either x and y are in the same cluster or they are not. If they are, only those σ such that $\sigma(x) = \sigma(y)$ can contribute to the last sum. Hence, the sum over σ is equal to $\frac{1}{2}$. If x and y are not connected by ω , the sum $\sum_{\sigma} \mathbb{S}_p(\sigma | \omega) \mathbb{1}_{\sigma(x) = \sigma(y)}$ is equal to $\frac{1}{2}$, thanks to Proposition 9.7 and the fact that exactly one half of the configurations σ compatible with ω assign the same spin to x and y . Finally,

$$\begin{aligned} \mathbb{Q}_\beta(\sigma(x) = \sigma(y)) - \frac{1}{2} &= \sum_{\omega} \mathbb{R}_{p,2}(\omega) \frac{1}{2} \mathbb{1}_{x \leftrightarrow y} + \left(\frac{1}{2} - \frac{1}{2} \right) \mathbb{1}_{x \not\leftrightarrow y} \\ &= \frac{1}{2} \mathbb{R}_{p,2}(x \leftrightarrow y), \end{aligned}$$

as expected. □

Thus, we have related the problem of the magnetization in the Ising model to the problem of connectivity in the random cluster model.

10 Infinite volume limits

10.1 Boundary conditions

We have proved that the correlation between the spins located at two distinct sites in the Ising model on a finite graph can be expressed in terms of the probability that these two sites are connected in the random cluster model. The magnetization in the Ising model is defined by putting boundary conditions to the spin configurations, namely setting all spins located on the boundary of a box equal to 1. We would like to prove that this corresponds to certain boundary conditions for the random cluster model, namely *wired* boundary conditions.

Let us choose an integer $n \geq 1$. Recall the definition of the subgraph $\Lambda_n = (V_n, E_n)$ of \mathbb{L}^d . The configuration space for the random cluster model on this graph is thus $\Omega_n = \{0, 1\}^{E_n}$. We say that a vertex $x \in V_n$ is on the boundary of Λ_n if one of its components

at least has modulus n . We say that an edge is on the boundary of Λ_n , and we write $e \in \partial\Lambda_n$, if it joins two vertices which are on the boundary.

Let us define a restricted configuration space $\Omega_n^w = \{\omega \in \Omega : \forall e \in \partial\Lambda_n, \omega(e) = 1\}$. The random cluster model with wired boundary conditions is the probability measure $\mathbb{R}_{p,2}^w$ obtained by conditioning $\mathbb{R}_{p,2}$ on the restricted configuration space.

Proposition 10.1 *Let p and β be related by $p = 1 - e^{-2\beta}$. Then, for all $n \geq 1$,*

$$\mathbb{Q}_{\beta,n}^+(\sigma(0) = 1) - \frac{1}{2} = \frac{1}{2} \mathbb{R}_{p,2}^w(0 \leftrightarrow \partial\Lambda_n).$$

Proof. Let us define a probability measure $\mathbb{S}_p^{+,w}$ by conditioning \mathbb{S}_p on the set $\Sigma_n^+ \times \Omega_n^w$. Then all the results of Section 9 remain true when $\mathbb{Q}_{\beta,n}$, $\mathbb{R}_{p,2}$ and \mathbb{S}_p are replaced respectively by $\mathbb{Q}_{\beta,n}^+$, $\mathbb{R}_{p,2}^w$ and $\mathbb{S}_p^{+,w}$. We leave the details as an exercise. \square

It is also still true that the measure $\mathbb{R}_{p,2}^w$ satisfies the FKG condition, hence the FKG inequality. This allows us to prove that there exists an infinite-volume limit for this measure as the size of the graph Λ_n grows to infinity.

Exercise 10.2 *Prove that the inequality (4) holds, and is in fact a strict inequality.*

10.2 The random cluster model on \mathbb{L}^d

Let $E \subset \mathbb{E}^d$ be a finite set of edges. Let A be an event on the configuration space $\{0, 1\}^{\mathbb{E}^d}$. Assume that A is a finite cylinder set which depends only on the edges of E . For n large enough, E is contained in the box $[-n, n]^d$ and the event A can be identified with an event on the configuration space Ω_n . Thus, it is assigned a probability $\mathbb{R}_{p,2}^{w,n}(A)$ for all n large enough. It turns out that, when A is increasing, these probabilities have a limit as n tends to infinity.

Proposition 10.3 *Let A be an increasing finite cylinder set. Then the sequence $(\mathbb{R}_{p,2}^{w,n}(A))_{n \geq n_0}$, defined for n large enough, is non-increasing. Its limit is denoted by $\mathbb{R}_{p,2}^w(A)$.*

Proof. Any event B on Ω_n can be identified with an event on Ω_{n+1} and the following equality holds:

$$\mathbb{R}_{p,2}^{w,n}(B) = \mathbb{R}_{p,2}^{w,n+1}(B | \text{all edges of } \partial\Lambda_n \text{ are open}).$$

The event that all edges of $\partial\Lambda_n$ are open is an increasing event. Thus, if B is increasing, it follows from the FKG inequality that

$$\mathbb{R}_{p,2}^{w,n}(B) \geq \mathbb{R}_{p,2}^{w,n+1}(B).$$

\square

We will admit that there exists a unique probability measure on $\{0, 1\}^{\mathbb{E}^d}$ which assigns the probability $\mathbb{R}_{p,2}^w(A)$ to every increasing finite cylinder set, and we will denote it by $\mathbb{R}_{p,2}^w$.

Proposition 10.4 *The following convergence holds:*

$$\mathbb{R}_{p,2}^{w,n}(0 \leftrightarrow \partial\Lambda_n) \xrightarrow{n \rightarrow \infty} \mathbb{R}_{p,2}^w(0 \leftrightarrow \infty).$$

Proof. Let us choose $n' \leq n$. Then $\mathbb{R}_{p,2}^{w,n}(0 \leftrightarrow \partial\Lambda_n) \leq \mathbb{R}_{p,2}^{w,n}(0 \leftrightarrow \partial\Lambda_{n'})$. As n tends to infinity, $\mathbb{R}_{p,2}^{w,n}(0 \leftrightarrow \partial\Lambda_{n'})$ tends, by Proposition 10.3, to $\mathbb{R}_{p,2}^w(0 \leftrightarrow \partial\Lambda_{n'})$. As n' tends to infinity, $\mathbb{R}_{p,2}^w(0 \leftrightarrow \partial\Lambda_{n'})$ tends to $\mathbb{R}_{p,2}^w(0 \leftrightarrow \infty)$, so that

$$\limsup_{n \rightarrow \infty} \mathbb{R}_{p,2}^{w,n}(0 \leftrightarrow \partial\Lambda_n) \leq \mathbb{R}_{p,2}^w(0 \leftrightarrow \infty).$$

On the other hand, $\mathbb{R}_{p,2}^{w,n}(0 \leftrightarrow \partial\Lambda_n) \geq \mathbb{R}_{p,2}^w(0 \leftrightarrow \partial\Lambda_n)$ by the monotonicity assertion in Proposition 10.3. Since $\mathbb{R}_{p,2}^w(0 \leftrightarrow \partial\Lambda_n)$ converges to $\mathbb{R}_{p,2}^w(0 \leftrightarrow \infty)$, we get

$$\liminf_{n \rightarrow \infty} \mathbb{R}_{p,2}^{w,n}(0 \leftrightarrow \partial\Lambda_n) \geq \mathbb{R}_{p,2}^w(0 \leftrightarrow \infty).$$

This finishes the proof. □

As a consequence of Propositions 10.1 and 10.4, the magnetization of the Ising model satisfies

$$M(\beta) = \frac{1}{2} \mathbb{R}_{p,2}^w(0 \leftrightarrow \infty),$$

where as usual $p = 1 - e^{-2\beta}$.

Proposition 10.5 *There is a phase transition for the random cluster model in any dimension $d \geq 2$: the critical probability*

$$p_c^{\text{RCM}}(d) = \sup\{p \in [0, 1] : \mathbb{R}_{p,2}^w(0 \leftrightarrow \infty) = 0\}$$

belongs to the open interval $(0, 1)$.

Proof. For $p < p_c(d)$, we have, since $\mathbb{R}_{p,2}^w \leq_{\text{st}} \mathbb{P}_p$,

$$\mathbb{R}_{p,2}^w(0 \leftrightarrow \infty) \leq \mathbb{P}_p(0 \leftrightarrow \infty) = 0.$$

Then, for $p > \frac{p_c(d)}{2-p_c(d)}$, we have

$$\mathbb{R}_{p,2}^w(0 \leftrightarrow \infty) \geq \mathbb{P}_{p'}(0 \leftrightarrow \infty) > 0.$$

The two inequalities imply the result. □

Finally, we have proved the following result.

Theorem 10.6 *In any dimension $d \geq 2$, there exists a critical value $\beta_c \in (0, +\infty)$ such that the magnetization of the Ising model on \mathbb{L}^d satisfies $M(\beta) = 0$ for $\beta > \beta_c$ and $M(\beta) > 0$ for $\beta < \beta_c$.*

We have thus proved the existence of a phase transition in the Ising model, in any dimension greater than 2.

Exercise 10.7 Assume that $d = 2$. Using the fact that $p_c(2) = \frac{1}{2}$, prove that

$$0.34 \leq \frac{\log 2}{2} \leq \beta_c(2) \leq \frac{\log 3}{2} \leq 0.55.$$

The actual value of $\beta_c(2)$ is $\frac{1}{2} \log(1 + \sqrt{2}) \simeq 0.44$.

References

- [1] Geoffrey Grimmett. *Percolation*, volume 321 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, second edition, 1999.
- [2] Geoffrey Grimmett. *Interacting Particles System*. Lectures delivered at Cambridge University. Personal notes, 2001.
- [3] Geoffrey Grimmett. *The random-cluster model*, volume 333 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, 2006.
- [4] David Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.