Geoffrey Grimmett

Percolation

Second Edition
With 121 Figures

Springer

Geoffrey Grimmett
Statistical Laboratory
University of Cambridge
16 Mill Lane
Cambridge CB2 1SB
United Kingdom
e-mail: G.R.Grimmett@statslab.cam.ac.uk

The first edition was published in 1989 as a monograph under the same title

Cataloging-in-Publication Data applied for

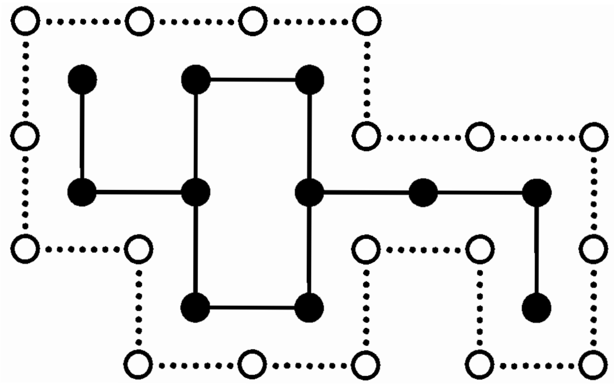
Die Deutsche Bibliothek – CIP-Einheitsaufnahme
Grimmett, Geoffrey: Percolation / Geoffrey Grimmett. – 2. ed. – Berlin; Heidelberg; New York; Barcelona; Hong Kong; London; Milan; Paris; Singapore; Tokyo: Springer, 1999
(Grundlehren der mathematischen Wissenschaften; 321)
ISBN 3-540-64902-6

Mathematics Subject Classification (1991): 60K35, 82B43

ISSN 0072-7830
ISBN 3-540-64902-6 Springer-Verlag Berlin Heidelberg New York

ISBN 3-540-96843-1 First edition Springer-Verlag New York Berlin Heidelberg

Cover design: MetaDesign plus GmbH, Berlin
Photocomposed from the author's TEX files.
SPIN: 10640056 41/3143-5 4 3 2 1 0 Printed on acid-free paper



Preface to the Second Edition

Quite apart from the fact that percolation theory has its origin in an honest applied problem, it is a source of fascinating problems of the best kind for which a mathematician can wish: problems which are easy to state with a minimum of preparation, but whose solutions are apparently difficult and require new methods. At the same time, many of the problems are of interest to or proposed by statistical physicists and not dreamed up merely to demonstrate ingenuity.

As a mathematical subject, percolation is a child of the 1950s. Following the presentation by Hammersley and Morton (1954) of a paper on Monte Carlo methods to the Royal Statistical Society, Simon Broadbent contributed the following to the discussion:

> "Another problem of excluded volume, that of the random maze, may be defined as follows: A square (in two dimensions) or cubic (in three) lattice consists of "cells" at the interstices joined by "paths" which are either open or closed, the probability that a randomly-chosen path is open being p. A "liquid" which cannot flow upwards or a "gas" which flows in all directions penetrates the open paths and fills a proportion $\lambda_r(p)$ of the cells at the rth level. The problem is to determine $\lambda_r(p)$ for a large lattice. Clearly it is a non-decreasing function of p and takes the values 0 at $p = 0$ and 1 at $p = 1$. Its value in the two-dimension case is not greater than in three dimensions.
>
> It appears likely from the solution of a simplified version of the problem that as $r \to \infty$ $\lambda_r(p)$ tends strictly monotonically to $\Lambda(p)$, a unique and stable proportion of cells occupied, independent of the way the liquid or gas is introduced into the first level. No analytical solution for a general case seems to be known."

This discussion led to a fruitful partnership between Broadbent and Hammersley, and resulted in their famous paper of 1957. The subsequent publications of Hammersley initiated the mathematical study of the subject.

Much progress has been made since, and many of the open problems of the last decades have been solved. With such solutions we have seen the evolution of new techniques and questions, and the consequent knowledge has shifted the ground under percolation. The mathematics of percolation is now fairly mature, although there are major questions which remain largely unanswered. Percolation technology has emerged as a cornerstone of the theory of disordered physical systems, and the methods of this book are now being applied and extended in a variety of important settings.

The quantity of literature related to percolation seems to grow hour by hour, mostly in the physics journals. It has become difficult to get to know the subject from scratch, and one of the principal purposes of this book is to remedy this. Percolation has developed a reputation for being hard as well as important. Nevertheless, it may be interesting to note that the level of mathematical preparation required to read this book is limited to some elementary probability theory and real analysis at the undergraduate level. Readers knowing a little advanced probability theory, ergodic theory, graph theory, or mathematical physics will not be disadvantaged, but neither will their knowledge aid directly their understanding of most of the hard steps.

This book is about the mathematics of percolation theory, with the emphasis upon presenting the shortest rigorous proofs of the main facts. I have made certain sacrifices in order to maximize the accessibility of the theory, and the major one has been to restrict myself almost entirely to the special case of bond percolation on the cubic lattice $\mathbb{Z}^d$. Thus there is only little discussion of such processes as continuum, mixed, inhomogeneous, long-range, first-passage, and oriented percolation. Nor have I spent much time or space on the relationship of percolation to statistical physics, infinite particle systems, disordered media, reliability theory, and so on. With the exception of the two final chapters, I have tried to stay reasonably close to core material of the sort which most graduate students in the area might aspire to know. No critical reader will agree entirely with my selection, and physicists may sometimes feel that my intuition is crooked.

Almost all the results and arguments of this book are valid for all bond and site percolation models, subject to minor changes only; the principal exceptions are those results of Chapter 11 which make use of the self-duality of bond percolation on the square lattice. I have no especially convincing reason for my decision to study bond percolation rather than the more general case of site percolation, but was swayed in this direction by historical reasons as well as the consequential easy access to the famous exact calculation of the critical probability of bond percolation on the square lattice. In addition, unlike the case of site models, it is easy to formulate a bond model having interactions which are long-range rather than merely nearest-neighbour. Such arguments indicate the scanty importance associated with this decision.

Here are a few words about the contents of this book. In the introductory Chapter 1 we prove the existence of a critical value p_c for the edge-probability p, marking the arrival on the scene of an infinite open cluster. The next chapter contains a general account of three basic techniques—the FKG and BK inequalities, and Russo's formula—together with certain other useful inequalities, some drawn from reliability theory. Chapter 3 contains a brief account of numerical equalities and inequalities for critical points, together with a general method for establishing strict inequalities. This is followed in Chapter 4 by material concerning the number of open clusters per vertex. Chapters 5 and 6 are devoted to subcritical percolation (with $p < p_c$). These chapters begin with the Menshikov and Aizenman–Barsky methods for identifying the critical point, and they continue with a systematic study of the subcritical phase. Chapters 7 and 8 are devoted to supercritical percolation (with $p > p_c$). They begin with an account of dynamic renormalization, the proof that percolation in slabs characterizes the supercritical phase, and a rigorous static renormalization argument; they continue with a deeper account of this phase. Chapter 9 contains a sketch of the physical approach to the critical phenomenon (when $p = p_c$), and includes an attempt to communicate to mathematicians the spirit of scaling theory and renormalization. Rigorous results are currently limited and are summarized in Chapter 10, where may be found the briefest sketch of the Hara–Slade mean field theory of critical percolation in high dimensions. Chapter 11 is devoted to percolation in two dimensions, where the technique of planar duality leads to the famous exact calculation that $p_c = \frac{1}{2}$ for bond percolation on $\mathbb{Z}^2$. The book terminates with two chapters of pencil sketches of related random processes, including continuum percolation, first-passage percolation, random electrical networks, fractal percolation, and the random-cluster model.

The first edition of this book was published in 1989. The second edition differs from the first through the reorganization of certain material, and through the inclusion of fundamental new material having substantial applications in broader contexts. In particular, the present volume includes accounts of strict inequalities between critical points, the relationship between percolation in slabs and in the whole space, the Burton–Keane proof of the uniqueness of the infinite cluster, the lace expansion and mean field theory, and numerous other results of significance. A full list of references is provided, together with pointers in the notes for each chapter.

A perennial charm of percolation is the beauty and apparent simplicity of its open problems. It has not been possible to do full justice here to work currently in progress on many such problems. The big challenge at the time of writing is to understand the proposal that critical percolation models in two dimensions are conformally invariant. Numerical experiments support this proposal, but rigorous verification is far from complete. While a full account of conformal invariance must await a later volume, at the ends of Chapters 9 and 11 may be found lists of references and a statement of Cardy's formula.

Most of the first edition of this book was written in draft form while I was visiting Cornell University for the spring semester of 1987, a visit assisted by a grant from the Fulbright Commission. It is a pleasure to acknowledge the assistance of Rick Durrett, Michael Fisher, Harry Kesten, Roberto Schonmann, and Frank Spitzer during this period. The manuscript was revised during the spring semester of 1988, which I spent at the University of Arizona at Tucson with financial support from the Center for the Study of Complex Systems and AFOSR contract no.

F49620-86-C-0130. One of the principal benefits of this visit was the opportunity for unbounded conversations with David Barsky and Chuck Newman. Rosine Bonay was responsible for the cover design and index of the first edition.

In writing the second edition, I have been aided by partial financial support from the Engineering and Physical Sciences Research Council under contract GR/L15425. I am grateful to Sarah Shea-Simonds for her help in preparing the TEXscript of this edition, and to Alexander Holroyd and Gordon Slade for reading and commenting on parts of it.

I make special acknowledgement to John Hammersley; not only did he oversee the early life of percolation, but also his unashamed love of a good problem has been an inspiration to many.

Unstinting in his help has been Harry Kesten. He read and commented in detail on much of the manuscript of the first edition, his suggestions for improvements being so numerous as to render individual acknowledgements difficult. Without his support the job would have taken much longer and been done rather worse, if at all.

G. R. G.
Cambridge
January 1999

Contents

Chapter 1

What is Percolation?

1.1 Modelling a Random Medium

Suppose we immerse a large porous stone in a bucket of water. What is the probability that the centre of the stone is wetted? In formulating a simple stochastic model for such a situation, Broadbent and Hammersley (1957) gave birth to the 'percolation model'. In two dimensions their model amounts to the following. Let $\mathbb{Z}^2$ be the plane square lattice and let p be a number satisfying $0 \le p \le 1$. We examine each edge of $\mathbb{Z}^2$ in turn, and declare this edge to be *open* with probability p and *closed* otherwise, independently of all other edges. The edges of $\mathbb{Z}^2$ represent the inner passageways of the stone, and the parameter p is the proportion of passages which are broad enough to allow water to pass along them. We think of the stone as being modelled by a large, finite subsection of $\mathbb{Z}^2$ (see Figure 1.1), perhaps those vertices and edges of $\mathbb{Z}^2$ contained in some specified connected subgraph of $\mathbb{Z}^2$. On immersion of the stone in water, a vertex x inside the stone is wetted if and only if there exists a path in $\mathbb{Z}^2$ from x to some vertex on the boundary of the stone, using open edges only. Percolation theory is concerned primarily with the existence of such 'open paths'.

If we delete the closed edges, we are left with a random subgraph of $\mathbb{Z}^2$; we shall study the structure of this subgraph, particularly with regard to the way in which this structure depends on the numerical value of p. It is not unreasonable to postulate that the fine structure of the interior passageways of the stone is on a scale which is negligible when compared with the overall size of the stone. In such circumstances, the probability that a vertex near the centre of the stone is wetted by water permeating into the stone from its surface will behave rather similarly to the probability that this vertex is the endvertex of an infinite path of open edges in $\mathbb{Z}^2$. That is to say, the large-scale penetration of the stone by water is related to the existence of infinite connected clusters of open edges.

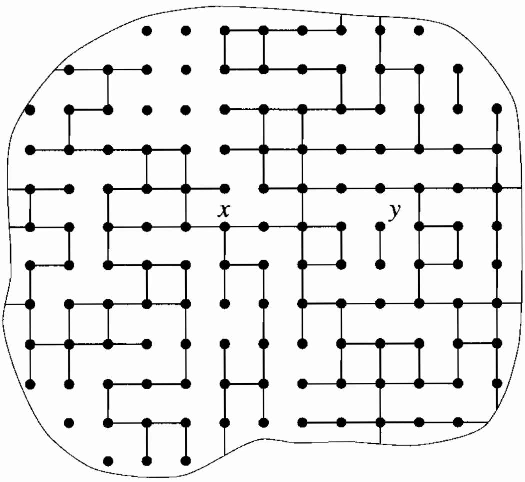


Figure 1.1. A sketch of the structure of a two-dimensional porous stone. The lines indicate the open edges; closed edges have been omitted. On immersion of the stone in water, vertex x will be wetted by the invasion of water, but vertex y will remain dry.

When can such infinite clusters exist? Simulations are handy indicators of the likely structure of the lattice, and Figure 1.2 contains such pictures for four different values of p. When $p = 0.25$, the connected clusters of open edges are isolated and rather small. As p increases, the sizes of clusters increase also, and there is a critical value of p at which there forms a cluster which pervades the entire picture. In loose terms, as we throw in more and more open edges, there comes a moment when large-scale connections are formed across the lattice. The pictures in Figure 1.2 are of course finite. If we were able to observe the whole of the infinite lattice $\mathbb{Z}^2$, we would see that all open clusters are finite when p is small, but that there exists an infinite open cluster for large values of p. In other words, there exists a critical value p_c for the edge-density p such that all open clusters are finite when $p < p_c$, but there exists an infinite open cluster when $p > p_c$ (such remarks should be interpreted 'with probability 1'). Drinkers of Pernod are familiar with this type of phenomenon—the transparence of a glass of Pernod is undisturbed by the addition of a small amount of water, but in the process of adding the water drop by drop, there arrives an instant at which the mixture becomes opaque.

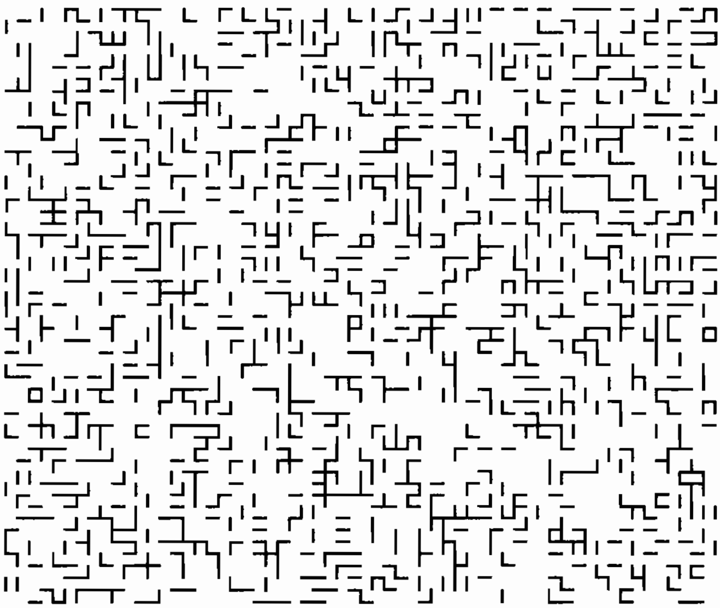
The occurrence of a 'critical phenomenon' is central to the appeal of percolation. In physical terms, we might say that the wetting of the stone is a 'surface effect' when the proportion p of open edges is small, and a 'volume effect' when p is large.

The above process is called 'bond percolation on the square lattice', and it is the most studied to date of all percolation processes. It is a very special process, largely because the square lattice has a certain property of self-duality which turns out to be extremely valuable. More generally, we begin with some periodic lattice in, say, d dimensions together with a number p satisfying $0 \le p \le 1$, and we declare each edge of the lattice to be open with probability p and closed otherwise. The resulting process is called a 'bond' model since the random blockages in the lattice are associated with the *edges*. Another type of percolation process is the 'site' percolation model, in which the *vertices* rather than the edges are declared to be open or closed at random, the closed vertices being thought of as junctions which are blocked to the passage of fluid. It is well known that every bond model may be reformulated as a site model on a different lattice, but that the converse is false (see Section 1.6). Thus site models are more general than bond models. They are illustrated in Figure 1.9.

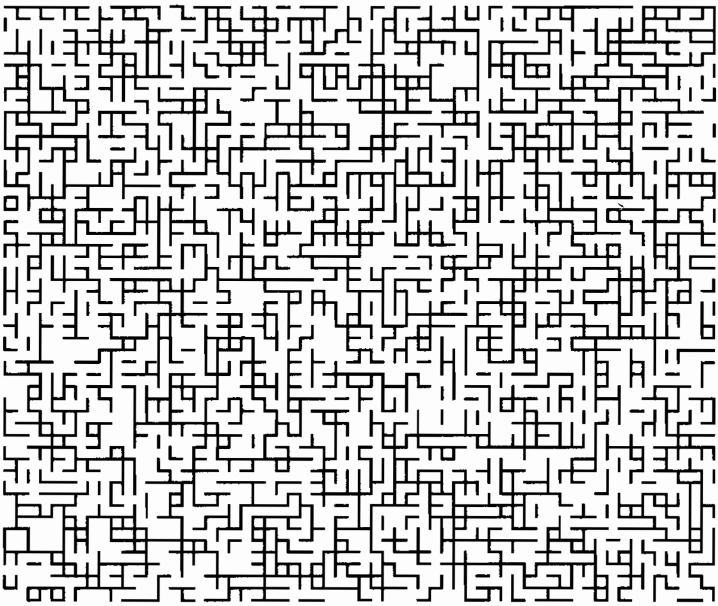
We may continue to generalize in several directions such as (i) 'mixed' models, in which both edges and vertices may be blocked, (ii) inhomogeneous models, in which different edges may have different probabilities of being open, (iii) long-range models, in which direct flow is possible between pairs of vertices which are very distant (in the above formulation, this may require a graph with large or even infinite vertex degrees), (iv) dependent percolation, in which the states of different edges are not independent, and so on. Mathematicians have a considerable talent in the art of generalization, and this has not been wasted on percolation theory. Such generalizations are often of considerable mathematical and physical interest; we shall however take the opposite route in this book. With few exceptions, we shall restrict ourselves to bond percolation on the d-dimensional cubic lattice $\mathbb{Z}^d$ where $d \ge 2$, and the main reason for this is as follows. As the level of generality rises, the accessibility of results in percolation theory is often diminished. Arguments which are relatively simple to explain in a special case can become concealed in morasses of geometrical and analytical detail when applied to some general model. This is not always the case, as illustrated by the proofs of exponential decay when $p < p_c$ (see Chapter 5) and of the uniqueness of the infinite open cluster when it exists (see Chapter 8). It is of course important to understand the limitations of an argument, but there may also be virtue in trying to describe something of the theory when stripped of peripheral detail. Bond percolation on $\mathbb{Z}^d$ is indeed a special case, but probably it exhibits the majority of properties expected of more general finite-range percolation-type models.

1.2 Why Percolation?

As a model for a disordered medium, percolation is one of the simplest, incorporating as it does a minimum of statistical dependence. Its attractions are manyfold. First, it is easy to formulate but not unrealistic in its qualitative predictions for random media. Secondly, for those with a greater interest in more complicated

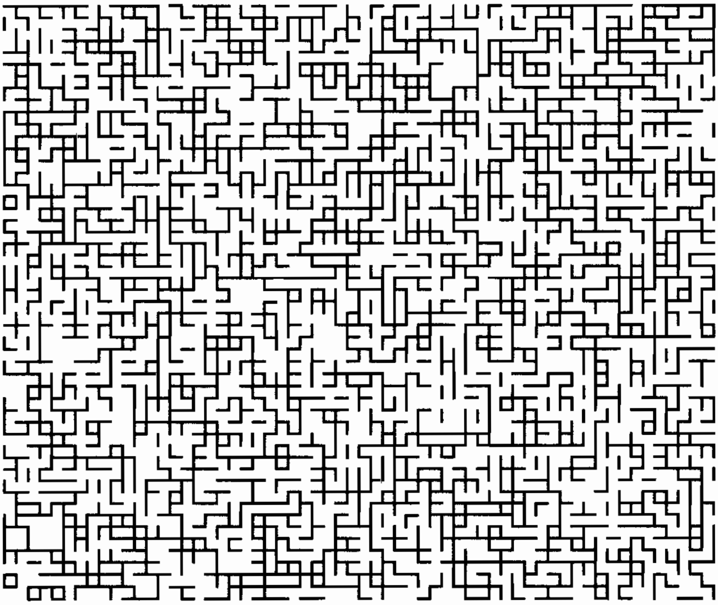


(a) $p = 0.25$

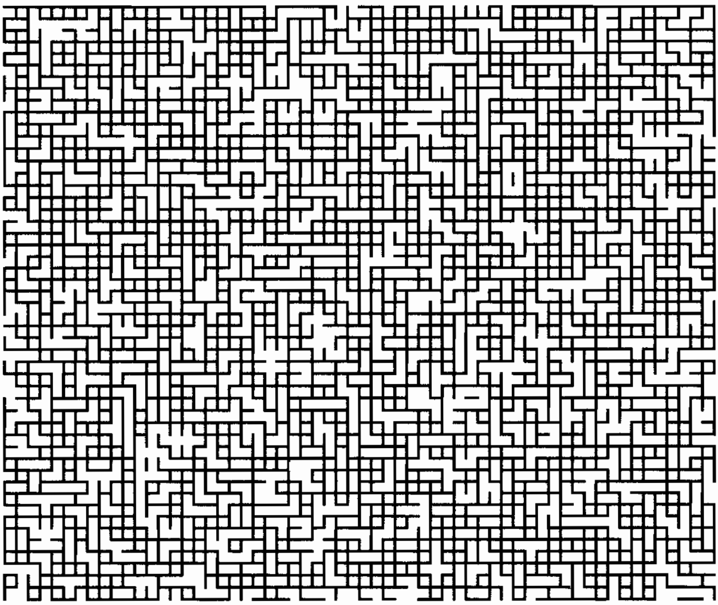


(b) $p = 0.49$

Figure 1.2. Realizations of bond percolation on a 50×60 section of the square lattice for four different values of p. The pictures have been created using the same sequence of pseudorandom numbers, with the result that each graph is a subgraph of the next. Readers with good



(c) $p = 0.51$



(d) $p = 0.75$

eyesight may care to check that there exist open paths joining the left to the right side when $p = 0.51$ but not when $p = 0.49$. The (random) value of p at which such paths appear for this realization is $0.5059\ldots$.

processes, it is a playground for developing mathematical techniques and insight. Thirdly, it is well endowed with beautiful conjectures which are easy to state but apparently rather hard to settle.

There is a fourth reason of significance. A great amount of effort has been invested in recent years towards an understanding of complex interacting random systems, including disordered media and other physical models. Such processes typically involve families of dependent random variables which are indexed by $\mathbb{Z}^d$ for some $d \geq 2$. To develop a full theory of such a system is often beyond the current methodology. Instead, one may sometimes obtain partial results by making a comparison with another process which is better understood. It is sometimes possible to make such a comparison with a percolation model. In this way, one may derive valuable results for the more complex system; these results may not be the best possible, but they may be compelling indicators of the directions to be pursued.

Here is an example. Consider a physical model having a parameter T called 'temperature'. It may be suspected that there exists a critical value T_c marking a phase transition. While this fact may itself be unproven, it may be possible to prove by comparison that the behaviour of the process for small T is qualitatively different from that for large T.

It has been claimed that percolation theory is a cornerstone of the theory of disordered media. As evidence to support this claim, we make brief reference to four types of disordered physical systems, emphasizing the role of percolation for each.

A. *Disordered electrical networks*. It may not be too difficult to calculate the effective electrical resistance of a block of either material A or of material B, but what is the effective resistance of a mixture of these two materials? If the mixture is disordered, it may be reasonable to assume that each component of the block is chosen at random to be of type A or of type B, independently of the types of all other components. The resulting effective resistance is a random variable whose distribution depends on the proportion p of components of type A. It seems to be difficult to say much of interest about the way in which this distribution depends on the numerical value of p. An extreme example arises when material B is a perfect insulator, and this is a case for which percolation comes to the fore. We illustrate this in a special example.

Let U_n be the square section $\{0, 1, \dots, n\} \times \{0, 1, \dots, n\}$ of the square lattice, and let S_n and T_n be the bottom and top sides of U_n,

$$S_n = \{(m, 0) : 0 \leq m \leq n\}, \qquad T_n = \{(m, n) : 0 \leq m \leq n\}.$$

We turn U_n into an electrical network as follows. We examine each edge of U_n in turn, and replace it by a wire of resistance 1 ohm with probability p, otherwise removing the connection entirely; this is done independently of all other edges. We now replace S_n and T_n by silver bars and we apply a potential difference between

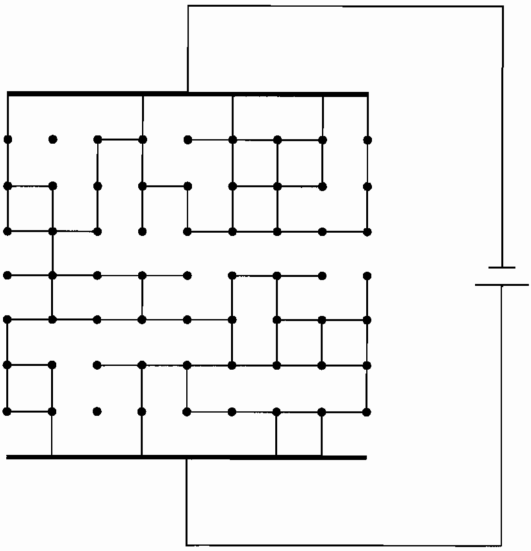


Figure 1.3. A realization of a random electrical network. Each remaining edge has unit resistance.

these bars; see Figure 1.3. What is the effective resistance R_n of the network? The value of R_n depends on the density and geometry of the set of edges having unit resistance, and such matters lie in the domain of percolation theory. We shall see in Section 13.2 that $R_n = \infty$ for all large n (almost surely) if $p < \frac{1}{2}$, whereas R_n is bounded uniformly away from 0 and ∞ if $p > \frac{1}{2}$.

B. *Ferromagnetism*. One of the most studied critical phenomena of theoretical physics is that of the ferromagnet. We position a lump of an appropriate metal in a magnetic field and we observe the way in which the magnetization of the metal varies according to imposed oscillations in the external field. Suppose that we increase the external field from 0 to some given value, and then decrease it back to 0. If the temperature is sufficiently large, the metal retains no residual magnetization, whereas at low temperatures the metal keeps some of its induced magnetization. There exists a critical value T_c of the temperature, called the Curie point, marking the borderline between the existence and non-existence of so called 'spontaneous magnetization'. A standard mathematical model for this phenomenon is the 'Ising model'. We give no definition of the Ising model here, but make instead some general remarks. In the Ising model on the lattice $\mathbb{Z}^d$, each vertex of $\mathbb{Z}^d$ may be in either of two states labelled 0 and 1. A configuration is an assignment $\omega = (\omega(x) : x \in \mathbb{Z}^d)$ of 0 or 1 to each vertex of the lattice. We consider probability measures on the set Ω of configurations taken in conjunction with some suitable σ-field of subsets of Ω; in particular, we are concerned with a class of measures having a type of 'spatial Markov' property: conditional on the states of all vertices outside any finite connected subgraph G of $\mathbb{Z}^d$, the states of vertices of G depend only on those of vertices in its 'external boundary'. Taken in conjunction with certain other conditions (positivity and translation invariance of the conditional probabilities, and positive correlation of increasing events), this property characterizes the class of measures of interest for the model. It turns out that there are two parameters which specify the conditional probabilities: the 'external magnetic field' h, and the strength J of interaction between neighbours. If $J = 0$, the states of different vertices are independent, and the process is equivalent to site percolation (see Section 1.6).

The relationship between the Ising model and bond percolation is rather strong. It turns out that they are linked via a type of 'generalized percolation' called the 'random-cluster model'. Through studying the random-cluster model, one obtains conclusions valid simultaneously for percolation and the Ising model. This important discovery was made around 1970 by Fortuin and Kasteleyn, and it has greatly influenced part of the current view of disordered physical systems. See Section 13.6 for a brief account of the random-cluster model.

C. *Epidemics and fires in orchards.* In an early review of percolation and related topics, Frisch and Hammersley (1963) proposed the use of percolation in modelling the spread of blight in a large orchard. The problem is as follows. Hypothetical trees are grown at the vertices of a square lattice. We suppose that there is a probability p that a healthy tree will be infected by a neighbouring blighted tree, where p is a known function of the distance between neighbouring trees. To prevent a single blighted tree from endangering a significant proportion of the whole orchard, it is necessary to choose the lattice spacing to be large enough that p is smaller than the critical probability of bond percolation on $\mathbb{Z}^2$.

In a forest fire, trees which are completely destroyed by fire cannot threaten their neighbours. Similarly, individuals who have recovered from some disease may gain protection from a recurrence of the disease. Such observations may be incorporated into a more complicated model which takes into account the passage of time. Suppose that each tree may be in any of three states: 1 (live and not on fire), 0 (burning), and -1 (burned). We suppose that the tree at vertex x burns for a random time T_x after catching fire, where $(T_x : x \in \mathbb{Z}^2)$ is a family of independent, identically distributed random variables. A burning tree emits sparks in the manner of a Poisson process with rate α, and each spark hits one of the neighbouring trees chosen at random; the spark sets fire to that tree so long as it is neither burned nor already on fire. At time 0, an arsonist sets light to the tree at the origin. It turns out that the set C of trees which are ultimately burned in the ensuing conflagration may be identified as the set of vertices reachable from the origin by open paths of a certain percolation-type process; this process differs from ordinary bond percolation in that the states of two different edges may be dependent if the edges have a vertex in common. See Section 13.5, as well as Cox and Durrett (1988), van den Berg, Grimmett, and Schinazi (1998), and the references therein.

D. *Wafer-scale integration.* In the manufacture of microchips, silicon wafers are engraved with copies of the required circuitry, these copies being laid out in a square grid. The wafer is then broken up into the individual chips, many of which are usually found to be faulty. After elimination of the faulty chips, the remaining non-defective chips are used to build processors. There are sound engineering and computing reasons for preferring to leave the wafer intact, making use of the non-defective chips in the positions in which they occurred in the wafer. Interconnections may be made between functioning units by using channels built between the rows and columns of the grid of chips. Such questions arise as the following:

(i) How long is the longest linear chain of functioning units which may be created using interconnections each of length not exceeding δ lattice units and laid in such a way that each channel of the wafer contains no more than two such interconnections?

(ii) Find the minimal interconnection length in a wiring pattern which creates a square grid of size $k \times k$ of functioning chips out of a wafer containing $n \times n$ units in all.

Greene and El Gamal (1984) answer such questions under the hypothesis that each chip is non-defective with probability p, independently of all other chips. Under this assumption, the set of functioning chips may be identified as the set of open vertices in a site percolation process on $\mathbb{Z}^2$, and thus the theory of percolation is relevant.

1.3 Bond Percolation

In this section we shall establish the basic definitions and notation of bond percolation on $\mathbb{Z}^d$. We begin with some graph theory. Throughout most of this book, the letter d stands for the dimension of the process; generally $d \geq 2$, but we assume for the moment only that $d \geq 1$. We write $\mathbb{Z} = \{\dots, -1, 0, 1, \dots\}$ for the set of all integers, and $\mathbb{Z}^d$ for the set of all vectors $x = (x_1, x_2, \dots, x_d)$ with integral coordinates. For $x \in \mathbb{Z}^d$ we generally write x_i for the ith coordinate of x. The (graph-theoretic) distance $\delta(x, y)$ from x to y is defined by

$$\delta(x, y) = \sum_{i=1}^{d} |x_i - y_i|, \tag{1.1}$$

and we write $|x|$ for the distance $\delta(0, x)$ from the origin to x. We shall sometimes have use for another distance function on $\mathbb{Z}^d$, and shall write

$$\|x\| = \max\{|x_i| : 1 \leq i \leq d\}, \tag{1.2}$$

noting that

$$\|x\| \leq |x| \leq d\|x\|.$$

We may turn $\mathbb{Z}^d$ into a graph, called the *d-dimensional cubic lattice*, by adding edges between all pairs x, y of points of $\mathbb{Z}^d$ with $\delta(x, y) = 1$. We denote this lattice by $\mathbb{L}^d$, and we write $\mathbb{Z}^d$ for the set of vertices of $\mathbb{L}^d$, and $\mathbb{E}^d$ for the set of its edges. In graph-theoretic terms, we write $\mathbb{L}^d = (\mathbb{Z}^d, \mathbb{E}^d)$. We shall often think of $\mathbb{L}^d$ as a graph embedded in $\mathbb{R}^d$, the edges being straight line segments between their endvertices. If $\delta(x, y) = 1$, we say that x and y are *adjacent*; in this case, we write $x \sim y$ and we represent the edge from x to y as $\langle x, y\rangle$. The edge e is *incident* to the vertex x if x is an endvertex of e. Letters such as u, v, w, x, y usually represent vertices, and letters such as e, f usually represent edges. We denote the origin of $\mathbb{Z}^d$ by 0.

Next we introduce probability. Let p and q satisfy $0 \leq p \leq 1$ and $p + q = 1$. We declare each edge of $\mathbb{L}^d$ to be *open* with probability p and *closed* otherwise, independently of all other edges. More formally, we consider the following probability space. As sample space we take $\Omega = \prod_{e\in\mathbb{E}^d}\{0, 1\}$, points of which are represented as $\omega = (\omega(e) : e \in \mathbb{E}^d)$ and called *configurations*; the value $\omega(e) = 0$ corresponds to e being closed, and $\omega(e) = 1$ corresponds to e being open. We take $\mathcal{F}$ to be the σ-field of subsets of Ω generated by the finite-dimensional cylinders. Finally, we take product measure with density p on $(\Omega, \mathcal{F})$; this is the measure

$$P_p = \prod_{e\in\mathbb{E}^d} \mu_e$$

where μ_e is Bernoulli measure on $\{0, 1\}$, given by

$$\mu_e\big(\omega(e) = 0\big) = q, \quad \mu_e\big(\omega(e) = 1\big) = p.$$

We write P_p for this product measure, and E_p for the corresponding expectation operator. We shall occasionally need a more general construction in which different edges may have different probabilities of being open. Such a construction begins with a family $\mathbf{p} = (p(e) : e \in \mathbb{E}^d)$ with $0 \leq p(e) \leq 1$ for all e. The appropriate probability space is now $(\Omega, \mathcal{F}, P_{\mathbf{p}})$ where $P_{\mathbf{p}} = \prod_{e\in\mathbb{E}^d} \mu_e$ and

$$\mu_e\big(\omega(e) = 0\big) = 1 - p(e), \quad \mu_e\big(\omega(e) = 1\big) = p(e)$$

for each e.

We write $\overline{A}$ (or occasionally A^{c}) for the complement of an event A, and I_A for the indicator function of A:

$$I_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A, \\ 0 & \text{if } \omega \notin A. \end{cases}$$

The expression $E_p(X; A)$ denotes the mean of X on the event A; that is to say, $E_p(X; A) = E_p(XI_A)$.

The following notation will be of value later. Let f be an edge of $\mathbb{L}^d$. We write P_p^f for Bernoulli product measure on $\prod_{e:e\neq f}\{0, 1\}$, the set of configurations of all edges of the lattice other than f. We think of P_p^f as being the measure associated with percolation on $\mathbb{L}^d$ with the edge f deleted.

There is a natural partial order on the set Ω of configurations, given by $\omega_1 \leq \omega_2$ if and only if $\omega_1(e) \leq \omega_2(e)$ for all $e \in \mathbb{E}^d$.

There is a one–one correspondence between Ω and the set of subsets of $\mathbb{E}^d$. For $\omega \in \Omega$, we define

$$K(\omega) = \{e \in \mathbb{E}^d : \omega(e) = 1\}; \tag{1.3}$$

thus $K(\omega)$ is the set of open edges of the lattice when the configuration is ω. Clearly, $\omega_1 \leq \omega_2$ if and only if $K(\omega_1) \subseteq K(\omega_2)$.

The following device can be useful. Suppose that $(X(e) : e \in \mathbb{E}^d)$ is a family of independent random variables indexed by the edge set $\mathbb{E}^d$, where each $X(e)$ is uniformly distributed on $[0, 1]$. We may couple together *all* bond percolation processes on $\mathbb{L}^d$ as p ranges over the interval $[0, 1]$ in the following way. Let p satisfy $0 \leq p \leq 1$ and define η_p $(\in \Omega)$ by

$$\eta_p(e) = \begin{cases} 1 & \text{if } X(e) < p, \\ 0 & \text{if } X(e) \geq p. \end{cases} \tag{1.4}$$

We say that the edge e is *p-open* if $\eta_p(e) = 1$. The random vector η_p has independent components and marginal distributions given by

$$P\big(\eta_p(e) = 0\big) = 1 - p, \quad P\big(\eta_p(e) = 1\big) = p.$$

We may think of η_p as being the random outcome of the bond percolation process on $\mathbb{L}^d$ with edge-probability p. It is clear that $\eta_{p_1} \leq \eta_{p_2}$ whenever $p_1 \leq p_2$, which is to say that we may couple together the two percolation processes with edge-probabilities p_1 and p_2 in such a way that the set of open edges of the first process is a subset of the set of open edges of the second. More generally, as p increases from 0 to 1, the configuration η_p runs through typical configurations of percolation processes with all edge-probabilities.

A *path* of $\mathbb{L}^d$ is an alternating sequence $x_0, e_0, x_1, e_1, \dots, e_{n-1}, x_n$ of distinct vertices x_i and edges $e_i = \langle x_i, x_{i+1}\rangle$; such a path has *length n* and is said to connect x_0 to x_n. A *circuit* of $\mathbb{L}^d$ is an alternating sequence $x_0, e_0, x_1, e_1, \dots, e_{n-1}, x_n, e_n, x_0$ of vertices and edges such that $x_0, e_0, \dots, e_{n-1}, x_n$ is a path and $e_n = \langle x_n, x_0\rangle$; such a circuit has length $n + 1$. We call a path or circuit *open* if all of its edges are open, and *closed* if all of its edges are closed. Two subgraphs of $\mathbb{L}^d$ are called *edge-disjoint* if they have no edges in common, and *disjoint* if they have neither edges nor vertices in common.

Consider the random subgraph of $\mathbb{L}^d$ containing the vertex set $\mathbb{Z}^d$ and the open edges only. The connected components of this graph are called *open clusters*. We write $C(x)$ for the open cluster containing the vertex x, and we call $C(x)$ the *open cluster at* x. The vertex set of $C(x)$ is the set of all vertices of the lattice which are connected to x by open paths, and the edges of $C(x)$ are the open edges of $\mathbb{L}^d$ which join pairs of such vertices. By the translation invariance of the lattice and of the probability measure P_p, the distribution of $C(x)$ is independent of the choice of x. The open cluster $C(0)$ at the origin is therefore typical of such clusters, and we represent this cluster by the single letter C. Occasionally we shall use the term $C(x)$ to represent the *set of vertices* joined to x by open paths, rather than the graph of this open cluster. We shall be interested in the size of $C(x)$, and we denote by $|C(x)|$ the number of vertices in $C(x)$. We note that $C(x) = \{x\}$ whenever x is incident to no open edge.

If A and B are sets of vertices of $\mathbb{L}^d$, we shall write '$A \leftrightarrow B$' if there exists an open path joining some vertex in A to some vertex in B; if $A \cap B \neq \varnothing$ then $A \leftrightarrow B$ trivially. Thus, for example, $C(x) = \{y \in \mathbb{Z}^d : x \leftrightarrow y\}$. We shall write '$A \nleftrightarrow B$' if there exists no open path from any vertex of A to any vertex of B, and '$A \leftrightarrow B$ off D' if there exists an open path joining some vertex in A to some vertex in B which uses no vertex in the set D.

If A is a set of vertices of the lattice, we write ∂A for the *surface* of A, being the set of vertices in A which are adjacent to some vertex not in A.

Our notation for boxes is the following. A *box* is a subset of $\mathbb{Z}^d$ of the form $B(a, b) = \{x \in \mathbb{Z}^d : a_i \leq x_i \leq b_i \text{ for all } i\}$, where a and b lie in $\mathbb{Z}^d$; we sometimes write

$$B(a, b) = \prod_{i=1}^{d} [a_i, b_i]$$

as a convenient shorthand. We often think of $B(a, b)$ as a subgraph of the lattice $\mathbb{L}^d$ suitably endowed with the edges which it inherits from the lattice. We denote by $B(n)$ the box with side-length $2n$ and centre at the origin:

(1.5) $$B(n) = [-n, n]^d = \{x \in \mathbb{Z}^d : \|x\| \leq n\}.$$

We may turn $B(n)$ into a graph by adding the edges which it inherits from $\mathbb{L}^d$. If x is a vertex of the lattice, we write $B(n, x)$ for the box $x + B(n)$ having side-length $2n$ and centre at x.

We write $\lfloor a \rfloor$ and $\lceil a \rceil$ for the integer part of the real number a, and the least integer not less than a, respectively. If $(a_n : n \geq 1)$ and $(b_n : n \geq 1)$ are sequences of real numbers, we write $a_n \sim b_n$ if $a_n/b_n \to 1$ as $n \to \infty$, and $a_n \approx b_n$ if $\log a_n / \log b_n \to 1$ as $n \to \infty$. Similarly, we write $f(p) \sim g(p)$ (respectively $f(p) \approx g(p)$) as $p \to \pi$ if $f(p)/g(p) \to 1$ (respectively $\log f(p)/\log g(p) \to 1$) as $p \to \pi$. Finally, we write $f(p) \asymp g(p)$ as $p \to \pi$ if $f(p)/g(p)$ is bounded away from 0 and ∞ on a neighbourhood of π.

1.4 The Critical Phenomenon

A principal quantity of interest is the *percolation probability* $\theta(p)$, being the probability that a given vertex belongs to an infinite open cluster. By the translation invariance of the lattice and probability measure, we lose no generality by taking this vertex to be the origin, and thus we define

$$\theta(p) = P_p(|C| = \infty). \tag{1.6}$$

Alternatively, we may write

$$\theta(p) = 1 - \sum_{n=1}^{\infty} P_p(|C| = n). \tag{1.7}$$

It is easy to see that $|C| = \infty$ if and only if there exists an infinite sequence $x_0, x_1, x_2, \ldots$ of distinct vertices such that $x_0 = 0$, $x_i \sim x_{i+1}$, and $\langle x_i, x_{i+1} \rangle$ is open for all i. Clearly θ is a non-decreasing function of p with $\theta(0) = 0$ and $\theta(1) = 1$. (Probably the most transparent proof of this monotonicity makes use of the coupling introduced around (1.4). See also Section 2.1.)

It is fundamental to percolation theory that there exists a critical value $p_c = p_c(d)$ of p such that

$$\theta(p) \begin{cases} = 0 & \text{if } p < p_c, \\ > 0 & \text{if } p > p_c; \end{cases}$$

$p_c(d)$ is called the *critical probability* and is defined formally by

$$p_c(d) = \sup\{p : \theta(p) = 0\}. \tag{1.8}$$

The case of one dimension is of no interest since, if $p < 1$, there exist infinitely many closed edges of $\mathbb{L}^1$ to the left and to the right of the origin almost surely, implying that $\theta(p) = 0$ if $p < 1$; thus $p_c(1) = 1$. The situation is quite different in two and more dimensions, and we shall see in Theorem (1.10) that $0 < p_c(d) < 1$ if $d \geq 2$. *We shall assume henceforth that, in the absence of an indication to the contrary, d is at least* 2. See Figure 1.4 for a sketch of the function θ.

The d-dimensional lattice $\mathbb{L}^d$ may be embedded in $\mathbb{L}^{d+1}$ in a natural way as the projection of $\mathbb{L}^{d+1}$ onto the subspace generated by the first d coordinates; with this embedding, the origin of $\mathbb{L}^{d+1}$ belongs to an infinite open cluster for a particular value of p whenever it belongs to an infinite open cluster of the sublattice $\mathbb{L}^d$. Thus $\theta(p) = \theta_d(p)$ is non-decreasing in d, which implies that

$$p_c(d+1) \leq p_c(d) \qquad \text{for } d \geq 1. \tag{1.9}$$

It is not very difficult to show that strict inequality is valid here, in that $p_c(d+1) < p_c(d)$ for all $d \geq 1$; see Section 3.3 and the notes for the current section.

The following theorem amounts to the statement that there exists a non-trivial critical phenomenon in dimensions two and more.

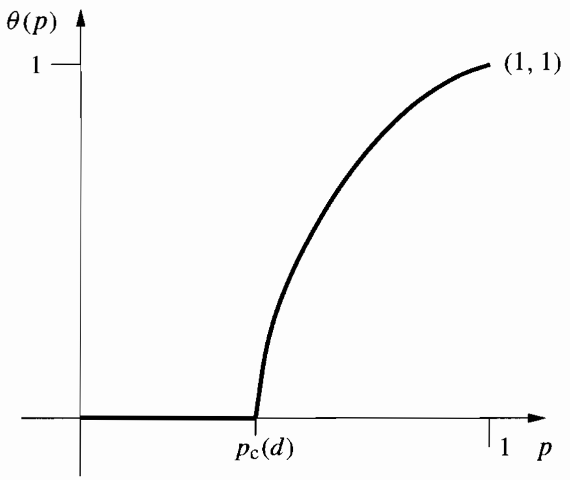


Figure 1.4. It is believed that the percolation probability $\theta(p)$ behaves roughly as indicated. It is known, for example, that θ is a continuous function of p except possibly at the critical probability $p_c(d)$. The possibility of a jump discontinuity at $p_c(d)$ has not been ruled out when $3 \leq d < 19$.

(1.10) Theorem. *If* $d \geq 2$ *then* $0 < p_c(d) < 1$.

The nub of this theorem is that in two or more dimensions there are two phases of the process. In the *subcritical* phase when $p < p_c(d)$, every vertex is almost surely in a finite open cluster, so that all open clusters are almost surely finite. In the *supercritical* phase when $p > p_c(d)$, each vertex has a strictly positive probability of being in an infinite open cluster, so that there exists almost surely at least one infinite open cluster. These phases are now reasonably well understood, which is more than can be said about the intermediate *critical* percolation process with $p = p_c(d)$, to which we shall return in more detail in Chapter 9. We make more concrete the above remarks about the subcritical and supercritical phases.

(1.11) Theorem. *The probability* $\psi(p)$ *that there exists an infinite open cluster satisfies*

$$\psi(p) = \begin{cases} 0 & \text{if } \theta(p) = 0, \\ 1 & \text{if } \theta(p) > 0. \end{cases}$$

Suppose in particular that $p = p_c(d)$. There exists an infinite open cluster if and only if $\theta(p_c(d)) > 0$. It turns out that no infinite open cluster exists when either $d = 2$ or $d \geq 19$, but it is an open question to determine whether or not there exists such a cluster for general d (including the physically important case $d = 3$); it is expected that no such cluster exists. Theorem (1.11) is proved by an application of the zero–one law, and this tells us nothing about the actual *number* of infinite open clusters when $\theta(p) > 0$; we shall however see in Section 8.2 that the infinite open cluster is (almost surely) unique whenever it exists.

Before proving these two theorems, we mention some associated results and open problems. First, what is the numerical value of $p_c(d)$? We know only the values $p_c(1) = 1$ and $p_c(2) = \frac{1}{2}$. The latter value is far from trivial to show, and this was the prize which attracted many people to the field in the 1970s. It is highly unlikely that there exists a useful representation of $p_c(d)$ for any other value of d, although such values may be calculated with increasing degrees of accuracy with the aid of larger and faster computers. Exact values are known for the critical probabilities of certain other two-dimensional lattices (for example, $p_c = 2\sin(\pi/18)$ for bond percolation on the triangular lattice); see Sections 3.1 and 11.9. It is the case that the value of the critical probability depends on both the dimension and the lattice structure, in contrast to certain asymptotic properties of $\theta(p)$ when p is near p_c: it is thought that, when $p - p_c$ is small and positive, then $\theta(p)$ behaves in a way which depends, to a degree, on the dimension d alone and is independent of the particular lattice structure. We return to this point in the next section and in Chapter 9.

Secondly, it is not difficult to find non-trivial upper and lower bounds for $p_c(d)$ when $d \geq 2$. We shall see in the proof of Theorem (1.10) that

(1.12) $$\frac{1}{\lambda(2)} \leq p_c(2) \leq 1 - \frac{1}{\lambda(2)},$$

and more generally

(1.13) $$\frac{1}{\lambda(d)} \leq p_c(d) \qquad \text{for } d \geq 3;$$

here, $\lambda(d)$ is the *connective constant* of $\mathbb{L}^d$, given by

(1.14) $$\lambda(d) = \lim_{n\to\infty} \{\sigma(n)^{1/n}\},$$

where $\sigma(n)$ is the number of paths (or 'self-avoiding walks') of $\mathbb{L}^d$ having length n and beginning at the origin. The exact value of $\lambda(d)$ is unknown for $d \geq 2$, but it is obvious that $\lambda(d) \leq 2d - 1$; to see this, note that each new step in a self-avoiding walk has at most $2d - 1$ choices since it must avoid the current position, and therefore $\sigma(n) \leq 2d(2d-1)^{n-1}$.

Thirdly, how does $p_c(d)$ behave when d is large? Inequality (1.13) implies that $(2d-1)p_c(d) \geq 1$, and it is known further that $p_c(d) \sim (2d)^{-1}$ as $d \to \infty$. This amounts to saying that, for large d, bond percolation on $\mathbb{L}^d$ behaves similarly to bond percolation on a regular tree in which each vertex has $2d(1 + o(1))$ descendants.

Proof of Theorem (1.10) and Equation (1.12). The existence of a critical phenomenon was shown by Broadbent and Hammersley (1957) and Hammersley (1957a, 1959).

We saw in (1.9) that $p_c(d+1) \leq p_c(d)$, and it suffices therefore to show that $p_c(d) > 0$ for $d \geq 2$, and that $p_c(2) < 1$. We prove first that $p_c(d) > 0$

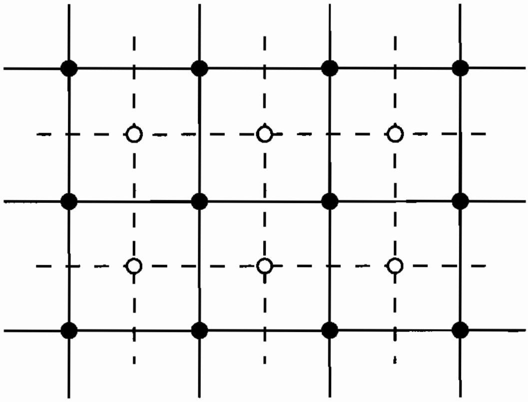


Figure 1.5. Part of the square lattice $\mathbb{L}^2$ together with its dual.

for $d \geq 2$. Consider bond percolation on $\mathbb{L}^d$ when $d \geq 2$. We shall show that $\theta(p) = 0$ whenever p is sufficiently close to 0. Let $\sigma(n)$ be the number of paths of $\mathbb{L}^d$ which have length n and which begin at the origin, and let $N(n)$ be the number of such paths which are open. Any such path is open with probability p^n, so that

$$E_p(N(n)) = p^n \sigma(n).$$

Now, if the origin belongs to an infinite open cluster then there exist open paths of all lengths beginning at the origin, so that

$$(1.15) \qquad \begin{aligned} \theta(p) &\leq P_p\big(N(n) \geq 1\big) \\ &\leq E_p(N(n)) = p^n \sigma(n) \qquad \text{for all } n. \end{aligned}$$

By the definition of the connective constant $\lambda(d)$ given at (1.14), we have that $\sigma(n) = \{\lambda(d) + \mathrm{o}(1)\}^n$ as $n \to \infty$; we substitute this into (1.15) to obtain

$$(1.16) \qquad \begin{aligned} \theta(p) &\leq \big\{p\lambda(d) + \mathrm{o}(1)\big\}^n \\ &\to 0 \qquad\qquad \text{if } p\lambda(d) < 1 \end{aligned}$$

as $n \to \infty$. Thus we have shown that $p_c(d) \geq \lambda(d)^{-1}$ where $\lambda(d) \leq 2d-1 < \infty$.

Secondly, we show that $p_c(2) < 1$, and we use an approach which is commonly called a 'Peierls argument' in honour of Rudolf Peierls and his 1936 article on the Ising model. Consider bond percolation on $\mathbb{L}^2$; we shall show that $\theta(p) > 0$ if p is sufficiently close to 1. It is here that planar duality is useful. Let G be a planar graph, drawn in the plane in such a way that edges intersect at vertices only. The *planar dual* of G is the graph obtained from G in the following way. We place a vertex in each face of G (including any infinite faces which may exist) and join two such vertices by an edge whenever the corresponding faces of G share a boundary edge in G. It is easy to see (especially with the aid of Figure 1.5) that the dual of $\mathbb{L}^2$

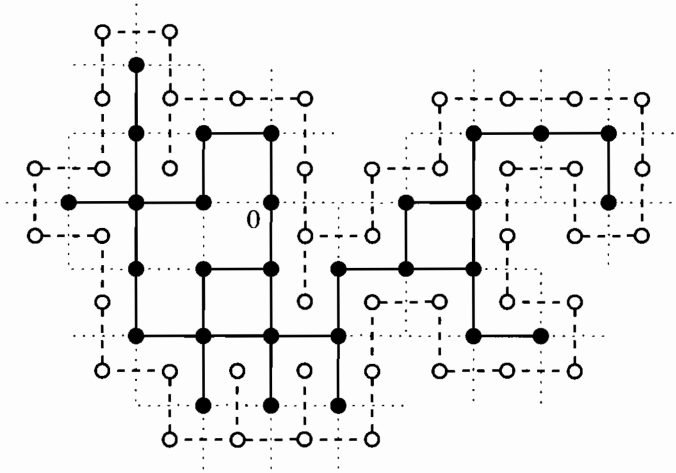


Figure 1.6. A finite open cluster at the origin, surrounded by a closed circuit in the dual lattice.

is isomorphic to $\mathbb{L}^2$; this *self*-duality is not in itself important at this stage, but will be crucial to our forthcoming proof in Chapter 11 that $p_c(2) = \frac{1}{2}$. For the sake of definiteness, we take as vertices of this dual lattice the set $\{x + (\frac{1}{2}, \frac{1}{2}) : x \in \mathbb{Z}^2\}$ and we join two such neighbouring vertices by a straight line segment of $\mathbb{R}^2$. There is a one–one correspondence between the edges of $\mathbb{L}^2$ and the edges of the dual, since each edge of $\mathbb{L}^2$ is crossed by a unique edge of the dual. We declare an edge of the dual to be open or closed depending respectively on whether it crosses an open or closed edge of $\mathbb{L}^2$. This assignment gives rise to a bond percolation process on the dual lattice with the same edge-probability p. We shall return to such matters in Chapter 11.

Suppose now that the open cluster at the origin of $\mathbb{L}^2$ is finite, and see Figure 1.6 for a sketch of the situation. We see that the origin is surrounded by a necklace of closed edges which are blocking off all possible routes from the origin to infinity. We may satisfy ourselves that the corresponding edges of the dual contain a closed circuit in the dual having the origin of $\mathbb{L}^2$ in its interior. This is best seen by inspecting Figure 1.6 again. It is somewhat tedious to formulate and prove such a statement with complete rigour, and we shall not do so here; see Kesten (1982, p. 386) for a more careful treatment. The converse holds similarly: if the origin lies in the interior of a closed circuit of the dual lattice, then the open cluster at the origin is finite. We summarize these remarks by saying that $|C| < \infty$ if and only if the origin of $\mathbb{L}^2$ lies in the interior of some closed circuit of the dual.

We now proceed as in the first part of the proof, by counting the number of such closed circuits in the dual. Let $\rho(n)$ be the number of circuits in the dual which have length n and which contain in their interiors the origin of $\mathbb{L}^2$. We may estimate $\rho(n)$ as follows. Each such circuit passes through some vertex of the form $(k + \frac{1}{2}, \frac{1}{2})$ for some k satisfying $0 \leq k < n$ since: first, it surrounds the origin and therefore passes through $(k+\frac{1}{2},\frac{1}{2})$ for some $k \geq 0$ and, secondly, it cannot pass through $(k+\frac{1}{2},\frac{1}{2})$ where $k \geq n$ since then it would have length at least $2n$. Thus such a circuit contains a self-avoiding walk of length $n-1$ starting from a vertex of the form $(k+\frac{1}{2},\frac{1}{2})$ where $0 \leq k < n$. The number of such self-avoiding walks is at most $n\sigma(n-1)$, giving that

$$\rho(n) \leq n\sigma(n-1). \tag{1.17}$$

Let γ be a circuit of the dual containing the origin of $\mathbb{L}^2$ in its interior, and let $M(n)$ be the number of such closed circuits having length n. By (1.17),

$$\begin{aligned}(1.18)\quad \sum_{\gamma} P_p(\gamma \text{ is closed}) &\leq \sum_{n=1}^{\infty} q^n n\sigma(n-1)\\ &= \sum_{n=1}^{\infty} qn\big\{q\lambda(2)+\mathrm{o}(1)\big\}^{n-1} && \text{as in (1.16)}\\ &< \infty && \text{if } q\lambda(2) < 1,\end{aligned}$$

where $q = 1-p$ and the summation is over all such γ. Furthermore,

$$\sum_{\gamma} P_p(\gamma \text{ is closed}) \to 0 \qquad \text{as } q = 1-p \downarrow 0,$$

so that we may find π satisfying $0 < \pi < 1$ such that

$$\sum_{\gamma} P_p(\gamma \text{ is closed}) \leq \tfrac{1}{2} \qquad \text{if } p > \pi.$$

It follows from the previous remarks that

$$\begin{aligned}P_p(|C|=\infty) &= P_p\big(M(n)=0 \text{ for all } n\big)\\ &= 1 - P_p\big(M(n) \geq 1 \text{ for some } n\big)\\ &\geq 1 - \sum_{\gamma} P_p(\gamma \text{ is closed})\\ &\geq \tfrac{1}{2} \qquad \text{if } p > \pi,\end{aligned}$$

giving that $p_c(2) \leq \pi$.

We need to work slightly harder in order to deduce that $p_c(2) \leq 1-\lambda(2)^{-1}$, as required for (1.12). The usual proof of this makes use of a certain correlation inequality known as the FKG inequality, which will be presented in Section 2.2. Rather than follow this usual route, we use a more elementary method which requires no extra technology. Let m be a positive integer. Let F_m be the event that there exists a closed dual circuit containing the box $B(m)$ in its interior, and let G_m

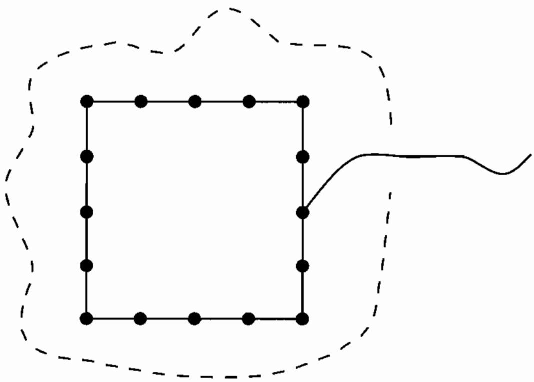


Figure 1.7. If there exists no closed dual circuit surrounding $B(m)$, then some vertex on the surface of $B(m)$ lies in an infinite open path.

be the event that all edges of $B(m)$ are open. These two events are independent, since they are defined in terms of disjoint sets of edges. Now, similarly to (1.18),

$$P_p(F_m) \le P_p\left(\sum_{n=4m}^{\infty} M(n) \ge 1\right) \le \sum_{n=4m}^{\infty} q^n n\sigma(n-1).$$

Much as before, if $q < \lambda(2)^{-1}$, we may find m such that $P_p(F_m) < \frac{1}{2}$, and we choose m accordingly. Assume now that G_m occurs but F_m does not. As indicated in Figure 1.7, the non-occurrence of F_m implies that some vertex of $B(m)$ lies in an infinite open path. Combined with the occurrence of G_m, this implies that $|C| = \infty$. Therefore, using the independence of F_m and G_m,

$$\theta(p) \ge P_p(\overline{F_m} \cap G_m) = P_p(\overline{F_m})P_p(G_m) \ge \tfrac{1}{2}P_p(G_m) > 0$$

if $q < \lambda(2)^{-1}$. □

Proof of Theorem (1.11). This is straightforward. First, we note that the event $\{\mathbb{L}^d$ contains an infinite open cluster$\}$ does not depend upon the states of any finite collection of edges. By the usual zero–one law (see, for example, Grimmett and Stirzaker (1992, p. 290)), ψ takes the values 0 and 1 only. If $\theta(p) = 0$ then

$$\psi(p) \le \sum_{x \in \mathbb{Z}^d} P_p\big(|C(x)| = \infty\big) = 0.$$

On the other hand, if $\theta(p) > 0$ then

$$\psi(p) \ge P_p(|C| = \infty) > 0$$

so that $\psi(p) = 1$ by the zero–one law, as required. □

1.5 The Main Questions

Consider bond percolation on $\mathbb{L}^d$ where $d \geq 2$. We are interested in the sizes and shapes of typical open clusters as the edge-probability p varies from 0 to 1, and we are particularly interested in large-scale phenomena such as the existence of infinite open clusters. We saw in the last section that 'macroscopic' quantities such as $\theta(p)$ and $\psi(p)$ have qualitatively different behaviour for small p than they have for large p. In addition to the probability that an open cluster is infinite, we may be interested in the mean size of an open cluster, and we write

$$\chi(p) = E_p|C| \tag{1.19}$$

for the mean number of vertices in the open cluster at the origin. Using the translation invariance of the process, we have that $\chi(p) = E_p|C(x)|$ for all vertices x. The functions θ and χ are two of the principal characters in percolation theory. We may express χ in terms of the distribution of $|C|$, just as we did for θ in (1.7):

$$\begin{aligned} \chi(p) &= \infty \cdot P_p(|C| = \infty) + \sum_{n=1}^{\infty} n P_p(|C| = n) \\ &= \infty \cdot \theta(p) + \sum_{n=1}^{\infty} n P_p(|C| = n), \end{aligned} \tag{1.20}$$

so that

$$\chi(p) = \infty \qquad \text{if } p > p_c. \tag{1.21}$$

The converse is not at all obvious: is it the case that $\chi(p) < \infty$ if $p < p_c$? We answer this question affirmatively in Chapter 5 (a sketch of the function χ appears in Figure 1.8). This indicates that the 'macroscopic' quantities θ and χ manifest critical behaviour at the same value of p. Indeed, most 'reasonable' macroscopic functions, such as θ and χ, are smooth functions of p except at the critical value p_c. It is commonly said that there exists a unique phase transition for percolation. More precisely, there are exactly two phases in the model—the subcritical phase $(p < p_c)$ and the supercritical phase $(p > p_c)$—together with the process *at* the critical point (when $p = p_c$). We shall study these phases in some detail in Chapters 5–10, but we present here a brief preview of some of the main results and open problems.

Subcritical phase. When $p < p_c$, all open clusters are finite almost surely. We shall see in Chapter 6 that $|C|$ has a tail which decreases exponentially, which is to say that there exists $\alpha(p)$ such that

$$P_p(|C| = n) \approx e^{-n\alpha(p)} \qquad \text{as } n \to \infty, \tag{1.22}$$

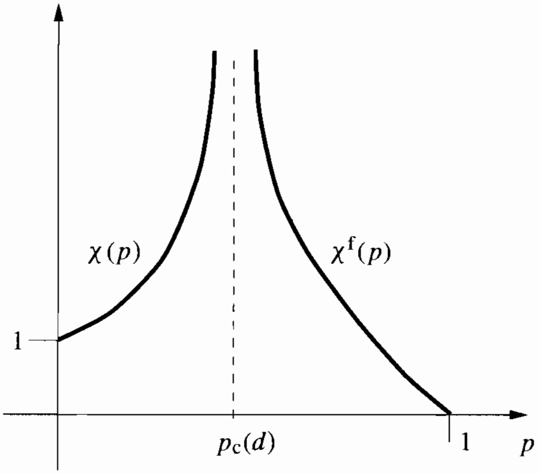


Figure 1.8. The left-hand curve is a sketch of the mean cluster size $\chi(p)$. The right-hand curve is a sketch of the mean size $\chi^{\mathrm{f}}(p)$ of a finite open cluster when $p > p_c$. Note that $\chi(p) = \chi^{\mathrm{f}}(p)$ if $p < p_c(d)$.

and $\alpha(p) > 0$ when $p < p_c$. It follows that $|C|$ has finite moments of all orders when $p < p_c$.

Supercritical phase. When $p > p_c$, there exist infinite open clusters almost surely, but how many? We shall see in Section 8.2 that the infinite open cluster is unique almost surely. If $|C| < \infty$ then how fast does the tail of $|C|$ decay? It is known that there exist $\beta_1(p)$ and $\beta_2(p)$, satisfying $0 < \beta_2(p) \le \beta_1(p) < \infty$, such that

$$(1.23) \qquad \exp\bigl(-\beta_1(p)n^{(d-1)/d}\bigr) \le P_p(|C| = n) \le \exp\bigl(-\beta_2(p)n^{(d-1)/d}\bigr) \quad \text{for all } n,$$

and it is believed that the limit

$$(1.24) \qquad \delta(p) = \lim_{n\to\infty}\Bigl\{-n^{-(d-1)/d}\log P_p(|C| = n)\Bigr\}$$

exists and is strictly positive when $p > p_c$. The basic reason for the power $n^{(d-1)/d}$ is that this is the order of the surface area of the sphere in $\mathbb{R}^d$ with volume n. The existence of the limit in (1.24) has been proved when $d = 2$ by Alexander, Chayes, and Chayes (1990), and when $d = 3$ by Cerf (1998b).

Since $\chi(p) = \infty$ when $p > p_c$, the function χ is of little interest in the supercritical phase. Instead, we concentrate on the related 'truncated' function given by

$$(1.25) \qquad \chi^{\mathrm{f}}(p) = E_p(|C|; |C| < \infty),$$

the mean of $|C|$ on the event that $|C| < \infty$. The function χ^{f} is probably not dissimilar in general form to the sketch in Figure 1.8. The superscript 'f' refers to the condition that C be finite.

At the critical point. It is hereabouts that we find major open problems. First, does there exist an infinite open cluster when $p = p_{\mathrm{c}}$? The answer is known to be negative when $d = 2$ or $d \geq 19$, and is generally believed to be negative for all $d \geq 2$. Assuming that $\theta(p_{\mathrm{c}}) = 0$, which is to say that there exists no infinite open cluster when $p = p_{\mathrm{c}}$, at what rate does $P_{p_{\mathrm{c}}}(|C| = n)$ decay? It is believed that

$$(1.26) \qquad P_{p_{\mathrm{c}}}(|C| \geq n) \approx n^{-1/\delta} \qquad \text{as } n \to \infty$$

for some $\delta = \delta(d) > 0$; the quantity δ is an example of a 'critical exponent'. Lower bounds for $P_p(|C| \geq n)$ of this general 'power' form are known for all dimensions $d \geq 2$, and also upper bounds when $d = 2$. Some have asked the provocative question "is it true that $\delta = \frac{91}{5}$ when $d = 2$, and $\delta = 2$ when $d \geq 6$?"; see Newman (1987a), for example.

Major progress has been made towards an understanding of critical percolation, but only under the assumption that d is sufficiently large. Currently the condition $d \geq 19$ suffices. When this holds, we know that $\theta(p_{\mathrm{c}}) = 0$, together with exact calculations of certain critical exponents.

Near the critical point. As p approaches p_{c} from above (or beneath), how do such quantities as $\theta(p)$ and $\chi(p)$ behave? It is commonly believed that they behave as powers of $|p - p_{\mathrm{c}}|$, and it is a major open problem of percolation to prove this. That is to say, we conjecture that the limits

$$(1.27) \qquad \gamma = -\lim_{p \uparrow p_{\mathrm{c}}} \frac{\log \chi(p)}{\log |p - p_{\mathrm{c}}|},$$

$$(1.28) \qquad \beta = \lim_{p \downarrow p_{\mathrm{c}}} \frac{\log \theta(p)}{\log (p - p_{\mathrm{c}})}$$

exist, and that the limit

$$(1.29) \qquad \delta^{-1} = -\lim_{n \to \infty} \frac{\log P_{p_{\mathrm{c}}}(|C| \geq n)}{\log n}$$

exists, in agreement with (1.26). The quantities γ, β, δ are called 'critical exponents'. There are physical reasons for believing the hypothesis of 'universality': the numerical values of critical exponents may depend only on the dimension d and not on the structure of the particular lattice. We return to such questions in Chapters 9 and 10, where we include a summary of progress towards answers to such questions. As remarked above, substantial progress has been made under the assumption that d is sufficiently large, currently that $d \geq 19$.

We close this section with a review of some of the principal characters in percolation. According to one method of counting, there are four such characters:

(a) the *percolation probability*

$$\theta(p) = P_p(|C| = \infty);$$

(b) the *mean size of the open cluster at the origin*

$$\chi(p) = E_p|C|;$$

(c) the *mean size of the finite open cluster at the origin*

$$\chi^{\mathrm{f}}(p) = E_p(|C|; |C| < \infty).$$

(d) The fourth such principal character is the *number of open clusters per vertex*, defined by

$$\kappa(p) = E_p(|C|^{-1}), \tag{1.30}$$

with the convention that $1/\infty = 0$. That is to say,

$$\kappa(p) = \sum_{n=1}^{\infty} \frac{1}{n} P_p(|C| = n).$$

We study the function κ in more detail in Chapter 4.

We observe that

$$\chi^{\mathrm{f}}(p) = \chi(p) \qquad \text{whenever } \theta(p) = 0. \tag{1.31}$$

There are many useful analogies between the percolation model and the Ising model, and we note that θ corresponds to magnetization, χ^{f} to susceptibility, and κ to free energy per vertex.

The quantities χ, χ^{f}, and κ are moments of the number of *vertices* in C. There are good reasons to define these quantities instead in terms of the number of *edges* of C, principally since such a definition would enable a unified approach to both bond and site percolation. For bond percolation on $\mathbb{L}^d$ it matters little which route we adopt, and we have chosen that which leads to fewest technical complications later.

1.6 Site Percolation

There are ways of impeding flow through a medium other than blocking the *edges*, and a natural alternative is to block the *vertices* instead. The corresponding model is termed 'site percolation', and it is defined as follows. We designate each *vertex* of the lattice $\mathbb{L}^d$ *open* with probability p, and *closed* otherwise; different vertices receive independent designations. A path is called *open* if all its vertices are open. The open cluster $C(x)$ at the vertex x is defined as the set of all vertices which may be attained by following open paths from x (if x is closed, then $C(x)$ is empty). As before, we write $C = C(0)$, and we define the *percolation probability* $\theta(p) = P_p(|C| = \infty)$, together with the *critical probability*

$$p_c = \sup\{p : \theta(p) = 0\}.$$

When we wish to emphasize the type of percolation model under study, we shall write θ^{site} or θ^{bond} (and p_c^{site} or p_c^{bond}) as appropriate.

Figure 1.9 contains four snapshots of site percolation on the square lattice for different values of p. The critical probability of this process is unknown, but is believed to be around 0.59; see Section 3.1.

Most arguments available for percolation models may be adapted to both bond and site models, and for that reason we pay only little attention to site percolation in this book. Indeed, there is a sense in which every bond model may be reformulated as a site model (on a different graph); the converse is false, and therefore site models are more general than bond models. We amplify this remark next. The *covering graph* (or *line graph*) of a graph G is the graph G_c defined as follows. To each edge of G there corresponds a distinct vertex of G_c, and two such vertices are deemed adjacent if and only if the corresponding edges of G share an endvertex. Suppose we are provided with a bond percolation process on G. We call a vertex of G_c *open* if and only if the corresponding edge of G is open. This induces a site percolation process on G_c. Furthermore, it is clear that every path of open edges in G corresponds to a path of open vertices in G_c (and vice versa). [We may note that there exist graphs on which the site model cannot be obtained from any bond model in the above way.]

Let us now consider an arbitrary infinite connected graph $G = (V, E)$. Let 0 denote a specified vertex of G which we call the 'origin'. We define $\theta^{\text{bond}}(p)$ (respectively $\theta^{\text{site}}(p)$) to be the probability that 0 lies in an infinite open cluster of G in a bond percolation (respectively site percolation) process on G having parameter p. Clearly $\theta^{\text{bond}}(p)$ and $\theta^{\text{site}}(p)$ are non-decreasing functions of p, and the bond and site critical probabilities are given by

$$p_c^{\text{bond}} = p_c^{\text{bond}}(G) = \sup\{p : \theta^{\text{bond}}(p) = 0\},$$
$$p_c^{\text{site}} = p_c^{\text{site}}(G) = \sup\{p : \theta^{\text{site}}(p) = 0\}.$$

We have from the above considerations that

$$p_c^{\text{bond}}(G) = p_c^{\text{site}}(G_c). \tag{1.32}$$

It is natural to ask whether there exists a relationship between the two critical points of a given graph G.

(1.33) Theorem. *Let $G = (V, E)$ be an infinite connected graph with countably many edges, origin 0, and maximum vertex degree Δ ($< \infty$). The critical probabilities of G satisfy*

$$\frac{1}{\Delta - 1} \le p_c^{\text{bond}} \le p_c^{\text{site}} \le 1 - \left(1 - p_c^{\text{bond}}\right)^{\Delta}. \tag{1.34}$$

One consequence of this theorem is that $p_c^{\text{bond}}(G) < 1$ if and only if $p_c^{\text{site}}(G) < 1$. The third inequality of (1.34) may be improved by replacing the exponent Δ by $\Delta - 1$, but we do not prove this here. Also, the strict inequality $p_c^{\text{bond}}(G) < p_c^{\text{site}}(G)$ is valid for a broad family of graphs G; see Section 3.4.

Proof. The first inequality of (1.34) follows by counting paths, as in (1.15)–(1.16). Therefore we turn immediately to the remaining two inequalities. In order to obtain these, we shall prove a certain stochastic inequality. Given two random subsets X, Y of V with associated expectation operator E, we write $X \le_{\text{st}} Y$, and say that X is *stochastically dominated* by Y, if

$$E(f(X)) \le E(f(Y))$$

for all bounded, measurable functions f satisfying $f(A) \le f(B)$ if $A \subseteq B \subseteq V$. A more systematic discussion of stochastic domination is provided in Section 7.4.

Let $C^{\text{bond}}(p)$ be a random subset of V having the law of the cluster of bond percolation at the origin; let $C^{\text{site}}(p)$ be a random subset having the law of the cluster of site percolation at the origin *conditional on* 0 *being an open vertex*. We claim that

$$C^{\text{site}}(p) \le_{\text{st}} C^{\text{bond}}(p) \tag{1.35}$$

and that

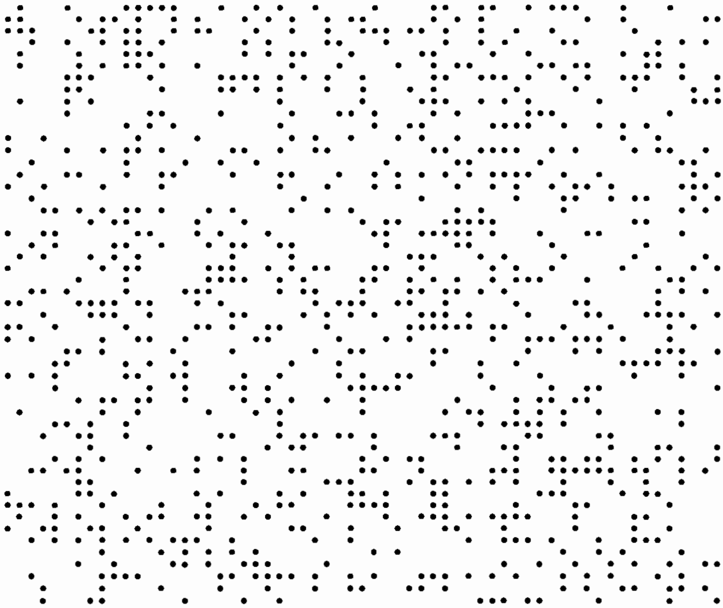
$$C^{\text{bond}}(p) \le_{\text{st}} C^{\text{site}}(p') \qquad \text{where } p' = 1 - (1 - p)^{\Delta}. \tag{1.36}$$

Since

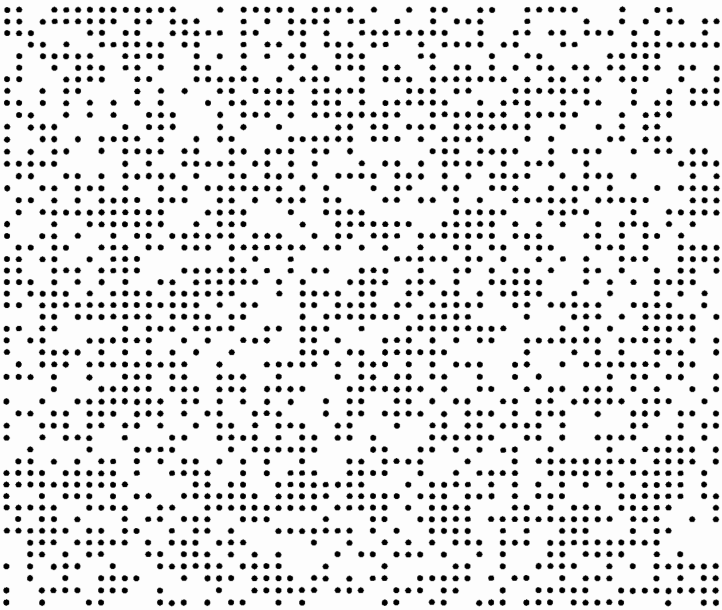
$$\theta^{\text{bond}}(p) = P_p\big(|C^{\text{bond}}(p)| = \infty\big),$$
$$p^{-1}\theta^{\text{site}}(p) = P_p\big(|C^{\text{site}}(p)| = \infty\big),$$

the remaining claims of (1.34) will follow from (1.35)–(1.36). Indeed, (1.35)–(1.36) imply that

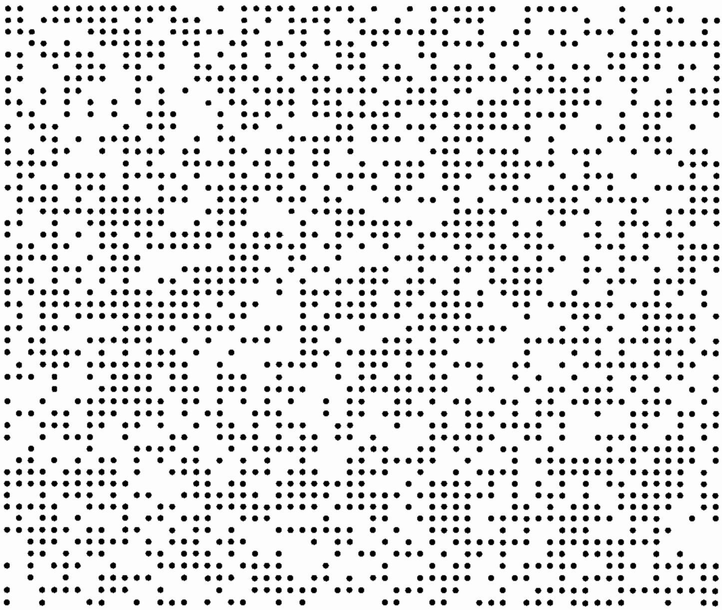
$$\frac{\theta^{\text{site}}(p)}{p} \le \theta^{\text{bond}}(p) \le \frac{\theta^{\text{site}}(p')}{p'} \qquad \text{where } p' = 1 - (1 - p)^{\Delta}, \tag{1.37}$$



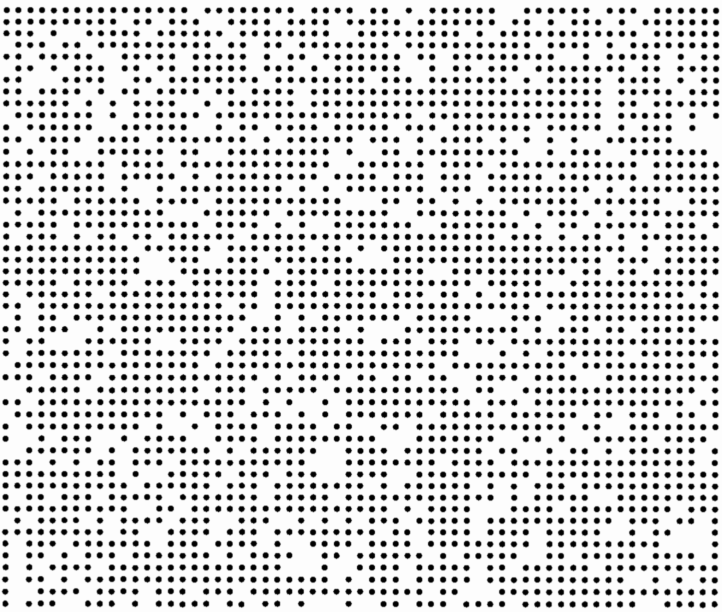
(a) $p = 0.3$



(b) $p = 0.58$



(c) $p = 0.60$



(d) $p = 0.80$

Figure 1.9. Realizations of site percolation on a 50×60 section of the square lattice for four different values of p. The critical probability of this process is believed to be near 0.59.

which is slightly stronger that the remaining parts of (1.34).

We construct appropriate couplings of the bond and site models in order to prove (1.35)–(1.36); this is a common technique when studying two processes simultaneously, and will be used later in this book. Let $\omega \in \{0, 1\}^E$ be a realization of a bond percolation process on $G = (V, E)$ having density p. We may build the cluster at the origin in the following standard manner. Let $e_1, e_2, \dots$ be a fixed ordering of E. At each stage k of the inductive construction, we shall have a pair (A_k, B_k) where $A_k \subseteq V$, $B_k \subseteq E$. Initially we set $A_0 = \{0\}$, $B_0 = \varnothing$. Having found (A_k, B_k) for some k, we define (A_{k+1}, B_{k+1}) as follows. We find the earliest edge e in the ordering of E having the following properties: $e \notin B_k$, and e is incident with exactly one vertex of A_k, say the vertex x. We now set

$$A_{k+1} = \begin{cases} A_k & \text{if } e \text{ is closed,} \\ A_k \cup \{y\} & \text{if } e \text{ is open,} \end{cases} \tag{1.38}$$

$$B_{k+1} = \begin{cases} B_k \cup \{e\} & \text{if } e \text{ is closed,} \\ B_k & \text{if } e \text{ is open,} \end{cases} \tag{1.39}$$

where $e = \langle x, y \rangle$. If no such edge e exists, we declare $(A_{k+1}, B_{k+1}) = (A_k, B_k)$. The sets A_k, B_k are non-decreasing, and the open cluster at the origin is given by $A_\infty = \lim_{k\to\infty} A_k$.

We now augment the above construction in the following way. We colour the vertex 0 *red*. Furthermore, on obtaining the edge e given above, we colour the vertex y *red* if e is open, and *black* otherwise. We specify that each vertex is coloured at most once in the construction, in the sense that any vertex y which is obtained at two or more stages is coloured in perpetuity according to the first colour it receives.

Let $A_\infty(\text{red})$ be the set of points connected to the origin by red paths of G (that is, by paths all of whose vertices are red). We make two claims concerning $A_\infty(\text{red})$:

(i) it is the case that $A_\infty(\text{red}) \subseteq A_\infty$, and all neighbours of vertices in $A_\infty(\text{red})$ which do not lie in $A_\infty(\text{red})$ are black;

(ii) $A_\infty(\text{red})$ has the same distribution as $C^{\text{site}}(p)$;

and inequality (1.35) follows immediately from these claims.

Claim (i) is straightforward. In order to be coloured red, a vertex is necessarily connected to the origin by a path of open edges. Furthermore, since all edges with exactly one endvertex in A_∞ are closed, all neighbours of $A_\infty(\text{red})$ which are not themselves coloured red are necessarily black.

We sketch an explanation of claim (ii). Whenever a vertex is coloured either red or black, it is coloured red with probability p, independently of all earlier colourings. This is not a full proof of (ii) but will satisfy many readers. More details are provided by Grimmett and Stacey (1998); see also the proof of Lemma (3.29).

The derivation of (1.36) is similar. We start with a directed version of G, namely the directed graph $\vec{G} = (V, \vec{E})$ obtained from G by replacing each edge

$e = \langle x, y \rangle$ by two directed edges, one in each direction, and denoted respectively by $[x, y\rangle$ and $[y, x\rangle$. We now let $\vec{\omega} \in \{0, 1\}^{\vec{E}}$ be a realization of an (oriented) bond percolation process on $\vec{G}$ having density p.

We colour the origin *green*. We colour a vertex x ($\neq 0$) *green* if at least one edge f of the form $[y, x\rangle$ satisfies $\vec{\omega}(f) = 1$; otherwise we colour x *black*. Then

$$(1.40) \qquad P_p(x \text{ is green}) = 1 - (1-p)^{\rho(x)} \leq 1 - (1-p)^{\Delta},$$

where $\rho(x)$ is the degree of x, and $\Delta = \max_x \rho(x)$.

We now build a copy A_∞ of $C^{\text{bond}}(p)$ more or less as described in (1.38)–(1.39). The only difference is that, on considering the edge $e = \langle x, y \rangle$ where $x \in A_k$, $y \notin A_k$, we declare e to be open for the purpose of (1.38)–(1.39) if and only if $\vec{\omega}([x, y\rangle) = 1$. Finally, we let $A_\infty(\text{green})$ be the set of points connected to the origin by green paths. It may be seen that $A_\infty(\text{green}) \supseteq A_\infty$. Furthermore, by (1.40), $A_\infty(\text{green})$ is stochastically dominated by $C^{\text{site}}(p')$ where $p' = 1 - (1-p)^{\Delta}$. Inequality (1.36) follows. □

1.7 Notes

Section 1.1. The origins of the mathematical theory of percolation may be found in the work of Flory (1941), Broadbent (1954), and Broadbent and Hammersley (1957). Hammersley (1983) and Grimmett (1999a) have described something of the history of the subject. We know of four books to date, the serious mathematical text of Kesten (1982), the gentle account by Efros (1986), and the books of Stauffer and Aharony (1991) and Hughes (1996). Of the many reviews, we mention Chayes and Chayes (1986a), Menshikov, Molchanov, and Sidorenko (1986), Aizenman (1987), Grimmett (1987b, 1997), Kesten (1987e), and Newman (1987a).

For general discussions of periodic lattices in d dimensions, see Grimmett (1978a, b), Godsil and McKay (1980), and Kesten (1982, Chapters 2, 3). Following Grimmett and Newman (1990), there has been serious study of percolation on graphs whose growth functions grow faster than any polynomial. Many such graphs arise as Cayley graphs of groups, and a systematic study of such systems has been initiated. See Benjamini and Schramm (1996), Benjamini, Lyons, Peres, and Schramm (1997), and the references therein.

The reader is referred to Chapter 12 for references appropriate to mixed, inhomogeneous, long-range, oriented, and first-passage percolation. The relationship between percolation and other models of statistical physics has been explored by Essam (1972, 1980); see also Section 13.6. Wierman (1987b) has studied 'high density' percolation in two dimensions; in this percolation-type process, one is interested only in those vertices which are incident to at least k open edges, for some specified value of k.

Section 1.3. The device for defining all percolation processes on the same probability space seems to have appeared for the first time in Hammersley (1963).

Section 1.4. The existence of a critical phenomenon was observed by Broadbent and Hammersley (1957) and Hammersley (1957a, 1959). Some people say that 'percolation occurs' when there exists an infinite open cluster.

Bond percolation on the line $\mathbb{Z}$ is essentially the problem of 'runs'; see Feller (1968). The proof that $p_c(2) = \frac{1}{2}$ was performed by Kesten (1980a), who built upon earlier arguments of Harris (1960), Russo (1978), and Seymour and Welsh (1978). Wierman (1981) adapted Kesten's proof to calculate exact values for critical probabilities for bond percolation on the triangular and hexagonal lattices.

Kesten (1982) proved that $p_c(3) < p_c(2)$; see also the related work of Menshikov, Molchanov, and Sidorenko (1986, Section 4). J. van den Berg (unpublished) and A. Frieze (unpublished) have pointed out a simple argument which leads to the strict inequality $p_c(d+1) < p_c(d)$ for $d \geq 1$. The argument amounts to the following for $d = 2$. Each edge of $\mathbb{L}^2$ may be thought of as the bottom of two infinite ladders of $\mathbb{L}^3$. By subdividing the 'vertical' edges in such ladders, we may construct disjoint ladders above different edges. By choosing the edge-probabilities for the subdivided edges with care, we arrive at the conclusion that $p_c(3) \leq 0.4798$, whereas $p_c(2) = \frac{1}{2}$.

Menshikov (1987a) has provided a more general argument (see also Zuev (1987)), showing that $p_c(\mathcal{L}_1) < p_c(\mathcal{L}_2)$ whenever $\mathcal{L}_2$ is a strict sublattice of $\mathcal{L}_1$ satisfying certain conditions; he was able to find non-trivial lower bounds for the difference $p_c(\mathcal{L}_2) - p_c(\mathcal{L}_1)$. His argument may be adapted to obtain a canonical approach to proving strict inequalities for general processes and enhancements thereof. See Aizenman and Grimmett (1991) and Sections 3.2–3.3 of the present book.

Kesten (1990) has proved that $p_c(d) \sim (2d)^{-1}$ as $d \to \infty$; see also Gordon (1991), Kesten (1991), and Kesten and Schonmann (1990). The detailed asymptotics of the forthcoming series expansion (3.2) for $p_c(d)$ were presented by Hara and Slade (1995). See also Hughes (1996).

The duality of planar lattices was explored by Hammersley (1959) and Harris (1960), and later by Fisher (1961). For an account of self-avoiding walks and the connective constant, see Hammersley (1957b); a recent account has been provided by Madras and Slade (1993).

Section 1.5. We defer the list of references until the appropriate forthcoming chapters.

Section 1.6. It was remarked by Fisher (1961) and Fisher and Essam (1961) that a bond model may be transformed into a site model; see also Kesten (1982, Chapter 3). We shall see at Theorem (2.8) that the definitions of $p_c^{\text{bond}}(G)$ and $p_c^{\text{site}}(G)$ are independent of the choice of origin, whenever the graph G is connected. Theorem (1.33) appeared in Grimmett (1997). The exponent Δ in the final inequality of (1.34) has been improved to $\Delta - 1$ by Grimmett and Stacey

(1998), where one may find also the strict inequality $p_c^{\text{bond}}(G) < p_c^{\text{site}}(G)$ for a broad class of graphs G including all finite-dimensional lattices in two or more dimensions. Further information concerning strict inequalities may be found in Chapter 3, and particularly in Section 3.4.

Chapter 2

Some Basic Techniques

2.1 Increasing Events

The probability $\theta(p)$, that the origin belongs to an infinite open cluster, is a non-decreasing function of p. This is intuitively obvious since an increase in the value of p leads generally to an increase in the number of open edges of $\mathbb{L}^d$, thereby increasing the number and lengths of open paths from the origin. Another way of putting this is to say that $\{|C| = \infty\}$ is an increasing event, in the sense that: if $\omega \in \{|C| = \infty\}$ then $\omega' \in \{|C| = \infty\}$ whenever $\omega \leq \omega'$. With such an example in mind we make the following definition. The event A in $\mathcal{F}$ is called *increasing* if $I_A(\omega) \leq I_A(\omega')$ whenever $\omega \leq \omega'$, where I_A is the indicator function of A. We call A *decreasing* if its complement $\overline{A}$ is increasing.

More generally, a random variable N on the measurable pair $(\Omega, \mathcal{F})$ is called *increasing* if $N(\omega) \leq N(\omega')$ whenever $\omega \leq \omega'$; N is called *decreasing* if $-N$ is increasing. Thus, an event A is increasing if and only if its indicator function is increasing.

As simple (and canonical) examples of increasing events and random variables, consider the event $A(x, y)$ that there exists an open path joining x to y, and the number $N(x, y)$ of different open paths from x to y.

The following result is intuitively clear.

(2.1) Theorem. *If N is an increasing random variable on $(\Omega, \mathcal{F})$, then*

(2.2) $$E_{p_1}(N) \leq E_{p_2}(N) \qquad \text{whenever } p_1 \leq p_2,$$

so long as these mean values exist. If A is an increasing event in $\mathcal{F}$, then

(2.3) $$P_{p_1}(A) \leq P_{p_2}(A) \qquad \text{whenever } p_1 \leq p_2.$$

There are good reasons for working with monotonic (that is, increasing or decreasing) events whenever possible. In doing calculations, we are often required to estimate such quantities as (i) the probability $P_p(A \cap B)$ that both A and B occur, in terms of the individual probabilities $P_p(A)$ and $P_p(B)$, or (ii) the rate of change of the probability $P_p(A)$ as p increases. For *monotonic* events A and B, there are general techniques for performing such estimates. We describe such techniques in the next three sections. The principal theorems of this chapter are the FKG and BK inequalities and Russo's formula. The reader is counselled to read at least Sections 2.2–2.4 before moving to the next chapter; some will choose to omit the formal proof of the BK inequality (Theorem (2.12)).

Before we prove Theorem (2.1), here is a general remark. The concept of monotonic events and the consequent techniques are rather general, and rely in no way upon the graph-theoretic structure of the percolation process. In the simple form in which we describe such techniques here, we require merely a probability space $(\Omega, \mathcal{F}, P)$ of the form $\Omega = \prod_{s \in S}\{0, 1\}$ where S is finite or countably infinite, $\mathcal{F}$ is the σ-field generated by the finite-dimensional cylinders of Ω, and P is a product measure on $(\Omega, \mathcal{F})$:

$$P = \prod_{s \in S} \mu_s$$

where μ_s is given by

$$\mu_s\big(\omega(s) = 0\big) = 1 - p(s), \quad \mu_s\big(\omega(s) = 1\big) = p(s)$$

for vectors $\omega = (\omega(s) : s \in S) \in \Omega$ and some specified collection $(p(s) : s \in S)$ of numbers from the interval $[0, 1]$. We have such a probability space to hand already, and thus we describe the basic techniques for dealing with increasing events in terms of this space. We may not always make explicit references to the underlying probability space, and so we adopt the following convention. Unless otherwise stated, the word 'event' shall refer to a member of $\mathcal{F}$, and the term 'random variable' shall refer to a measurable real-valued function on $(\Omega, \mathcal{F})$, where $(\Omega, \mathcal{F}, P_p)$ is the basic probability space of bond percolation on $\mathbb{L}^d$.

Proof of Theorem (2.1). Let the random variables $(X(e) : e \in \mathbb{E}^d)$ be independent and uniformly distributed on $[0, 1]$. As in Section 1.3, we write $\eta_p(e) = 1$ if $X(e) < p$ and $\eta_p(e) = 0$ otherwise. If $p_1 \le p_2$ then $\eta_{p_1} \le \eta_{p_2}$, giving that $N(\eta_{p_1}) \le N(\eta_{p_2})$ for any increasing random variable N on $(\Omega, \mathcal{F})$. We take expectations of this inequality to obtain (2.2). The second part of the theorem follows immediately by applying the first part to the increasing random variable $N = I_A$. □

2.2 The FKG Inequality

It is highly plausible that increasing events A and B are positively correlated, in the sense that $P_p(A \cap B) \geq P_p(A)P_p(B)$. For example, if we know that there exists an open path joining vertex u to vertex v, then it becomes more likely than before that there exists an open path joining vertex x to vertex y. Such a correlation inequality was first proved by Harris (1960). Since then, such inequalities have been reworked in a more general context than product measure on $(\Omega, \mathcal{F})$; by current convention, they are named after Fortuin, Kasteleyn, and Ginibre (1971).

(2.4) Theorem. FKG inequality.

(a) *If X and Y are increasing random variables such that $E_p(X^2) < \infty$ and $E_p(Y^2) < \infty$, then*

(2.5) $$E_p(XY) \geq E_p(X)E_p(Y).$$

(b) *If A and B are increasing events then*

(2.6) $$P_p(A \cap B) \geq P_p(A)P_p(B).$$

Similar inequalities are valid for decreasing random variables and events. For example, if X and Y are both decreasing then $-X$ and $-Y$ are increasing, giving that

$$E_p(XY) \geq E_p(X)E_p(Y)$$

so long as X and Y have finite second moments. Similarly, if X is increasing and Y is decreasing, then we may apply the FKG inequality to X and $-Y$ to find that

$$E_p(XY) \leq E_p(X)E_p(Y).$$

As an example of the FKG inequality in action, we describe the way in which it is normally applied. Let $\Pi_1, \Pi_2, \ldots, \Pi_k$ be families of paths in $\mathbb{L}^d$, and let A_i be the event that there exists some path in Π_i which is open. The A_i are increasing events, and therefore

$$P_p\left(\bigcap_{i=1}^{k} A_i\right) \geq P_p(A_1)P_p\left(\bigcap_{i=2}^{k} A_i\right),$$

since the intersection of increasing events is increasing. We iterate this to obtain

(2.7) $$P_p\left(\bigcap_{i=1}^{k} A_i\right) \geq \prod_{i=1}^{k} P_p(A_i).$$

Here is a concrete application of the FKG inequality. Let $G = (V, E)$ be an infinite connected graph with countably many edges, and consider a bond percolation process on G. For any vertex x, we write $\theta(p, x)$ for the probability that x lies in an infinite open cluster, and

$$p_c(x) = \sup\{p : \theta(p, x) = 0\}$$

for the associated critical probability. We have by the FKG inequality that

$$\theta(p, x) \geq P_p\big(\{x \leftrightarrow y\} \cap \{y \leftrightarrow \infty\}\big) \geq P_p(x \leftrightarrow y)\theta(p, y),$$

whence $p_c(x) \leq p_c(y)$. The latter inequality holds also with x and y interchanged. A similar argument is valid for site percolation, and we arrive at the following theorem.

(2.8) Theorem. *Let G be an infinite connected graph with countably many edges. The values of the bond critical probability $p_c^{\text{bond}}(x)$ and the site critical probability $p_c^{\text{site}}(x)$ are independent of the choice of initial vertex x.*

Proof of FKG inequality. We need only prove part (a), since part (b) follows by applying the first part to the indicator functions of A and B. First we prove part (a) for random variables X and Y which are defined in terms of the states of only finitely many edges; later we shall remove this restriction.

Suppose that X and Y are increasing random variables which depend only on the states of the edges $e_1, e_2, \dots, e_n$. We proceed by induction on n. First, suppose $n = 1$. Then X and Y are functions of the state $\omega(e_1)$ of e_1, which takes the values 0 and 1 with probabilities $1 - p$ and p, respectively. Now,

$$\big\{X(\omega_1) - X(\omega_2)\big\}\big\{Y(\omega_1) - Y(\omega_2)\big\} \geq 0$$

for all pairs ω_1, ω_2, each taking the value 0 or 1; this is trivial if $\omega_1 = \omega_2$, and follows from the monotonicity of X and Y otherwise. Thus

$$\begin{aligned} 0 \leq \sum_{\omega_1, \omega_2} & \big\{X(\omega_1) - X(\omega_2)\big\}\big\{Y(\omega_1) - Y(\omega_2)\big\} \\ & \qquad \times P_p\big(\omega(e_1) = \omega_1\big) P_p\big(\omega(e_1) = \omega_2\big) \\ = 2\big\{ & E_p(XY) - E_p(X)E_p(Y)\big\} \end{aligned}$$

as required. Suppose now that the result is valid for values of n satisfying $n < k$, and that X and Y are increasing functions of the states $\omega(e_1), \omega(e_2), \dots, \omega(e_k)$ of the edges $e_1, e_2, \dots, e_k$. Then

$$\begin{aligned} E_p(XY) &= E_p\Big(E_p\big(XY \,\big|\, \omega(e_1), \dots, \omega(e_{k-1})\big)\Big) \\ &\geq E_p\Big(E_p\big(X \,\big|\, \omega(e_1), \dots, \omega(e_{k-1})\big) E_p\big(Y \,\big|\, \omega(e_1), \dots, \omega(e_{k-1})\big)\Big), \end{aligned}$$

since, for given $\omega(e_1), \dots, \omega(e_{k-1})$, it is the case that X and Y are increasing in the single variable $\omega(e_k)$. Now, $E_p(X \mid \omega(e_1), \dots, \omega(e_{k-1}))$ is an increasing function of the states of $k - 1$ edges, as is the corresponding function of Y.

It follows from the induction hypothesis that the last mean value above is no smaller than the product of the means, whence

$$\begin{aligned} E_p(XY) \geq\ & E_p\Big(E_p\big(X \,\big|\, \omega(e_1), \dots, \omega(e_{k-1})\big)\Big) \\ & \qquad \times E_p\Big(E_p\big(Y \,\big|\, \omega(e_1), \dots, \omega(e_{k-1})\big)\Big) \\ =\ & E_p(X)E_p(Y). \end{aligned}$$

We now lift the condition that X and Y depend on the states of only finitely many edges. Suppose that X and Y are increasing random variables with finite second moments. Let $e_1, e_2, \dots$ be a (fixed) ordering of the edges of $\mathbb{L}^d$, and define

$$X_n = E_p\big(X \,\big|\, \omega(e_1), \dots, \omega(e_n)\big), \quad Y_n = E_p\big(Y \,\big|\, \omega(e_1), \dots, \omega(e_n)\big). \tag{2.9}$$

Now X_n and Y_n are increasing functions of the states of $e_1, e_2, \dots, e_n$, and therefore,

$$E_p(X_nY_n) \geq E_p(X_n)E_p(Y_n) \tag{2.10}$$

by the discussion above. As a consequence of the martingale convergence theorem (see Grimmett and Stirzaker (1992, p. 309)) we have that, as $n \to \infty$,

$$X_n \to X \quad \text{and} \quad Y_n \to Y \qquad P_p\text{-a.s. and in } L^2(P_p),$$

whence

$$E_p(X_n) \to E_p(X) \quad \text{and} \quad E_p(Y_n) \to E_p(Y) \qquad \text{as } n \to \infty. \tag{2.11}$$

Also, by the triangle and Cauchy–Schwarz inequalities,

$$\begin{aligned} E_p|X_nY_n - XY| &\leq E_p\big(|(X_n - X)Y_n| + |X(Y_n - Y)|\big) \\ &\leq \sqrt{E_p\big((X_n - X)^2\big)E_p(Y_n^2)} + \sqrt{E_p(X^2)E_p\big((Y_n - Y)^2\big)} \\ &\to 0 \qquad \text{as } n \to \infty, \end{aligned}$$

so that $E_p(X_nY_n) \to E_p(XY)$. We take the limit as $n \to \infty$ in (2.10) to obtain the result. □

2.3 The BK Inequality

The positive correlation of increasing events is only half of the story; in many situations the direction of the FKG inequality is useless for the problem, and so we seek some complementary correlation-type inequality. It is, of course, hopeless to expect some useful general upper bound on $P_p(A \cap B)$ for increasing events A and B. It turns out that the right way to approach the question is to replace the event $A \cap B$, that both A and B occur, by the event $A \circ B$, that A and B occur on disjoint edge sets of $\mathbb{L}^d$. We make precise this idea of 'disjoint occurrence' as follows.

Let $e_1, e_2, \dots, e_n$ be n distinct edges of $\mathbb{L}^d$, and let A and B be increasing events which depend on the vector $\omega = (\omega(e_1), \omega(e_2), \dots, \omega(e_n))$ of the states of these edges only. Each such ω is specified uniquely by the set $K(\omega) = \{e_i : \omega(e_i) = 1\}$ of edges with state 1. We define the event $A \circ B$ to be the set of all ω for which there exists a subset H of $K(\omega)$ such that ω', determined by $K(\omega') = H$, belongs to A, and ω'', determined by $K(\omega'') = K(\omega) \backslash H$, belongs to B. We say that $A \circ B$ is the event that A and B 'occur disjointly'. Thus $A \circ B$ is the set of configurations ω for which there exist disjoint sets of open edges with the property that the first such set guarantees the occurrence of A and the second guarantees the occurrence of B.

The following is the canonical example of disjoint occurrence. Let G be a finite subgraph of $\mathbb{L}^d$, and let $A_G(x, y)$ be the event that there exists an open path joining vertex x of G to vertex y of G, lying entirely within G. Then $A_G(u, v) \circ A_G(x, y)$ is the event that there exist two edge-disjoint open paths in G, the first joining u to v and the second joining x to y. Suppose now that we are given that $A_G(u, v)$ occurs, and we ask for the (conditional) probability of $A_G(u, v) \circ A_G(x, y)$. The conditioning on $A_G(u, v)$ amounts to knowing some information about the occurrence of open edges, but we are not allowed to use all such open edges in finding an open path from x to y disjoint from one of the open paths from u to v. Thus it is plausible that

$$P_p\big(A_G(u, v) \circ A_G(x, y) \,\big|\, A_G(u, v)\big) \le P_p\big(A_G(x, y)\big).$$

This is essentially the assertion of the BK inequality, proved by van den Berg and Kesten (1985).

We shall state the BK inequality in a slightly more general context than that described above. Let m be a positive integer, let $\Gamma = \prod_{i=1}^m \{0, 1\}$, $\mathcal{G}$ be the set of all subsets of Γ, and P be product measure on $(\Gamma, \mathcal{G})$ with density $p(i)$ on the ith coordinate of vectors in Γ; thus

$$P = \prod_{i=1}^{m} \mu_i,$$

where $\mu_i(0) = 1 - p(i)$ and $\mu_i(1) = p(i)$.

(2.12) Theorem. BK inequality. *If A and B are increasing events in $\mathcal{G}$, then*

$$P(A \circ B) \leq P(A)P(B). \tag{2.13}$$

We note that $A \circ B \subseteq A \cap B$ and that $A \circ B$ is increasing whenever A and B are increasing. The BK inequality may be generalized in the following direction. Let $A_1, A_2, \ldots, A_k$ be increasing events in $\mathcal{G}$. It is easily checked that the operation $\circ$ is associative, in that $A_i \circ (A_j \circ A_k) = (A_i \circ A_j) \circ A_k$, and repeated application of (2.13) yields

$$P(A_1 \circ A_2 \circ \cdots \circ A_k) \leq \prod_{i=1}^{k} P(A_i). \tag{2.14}$$

In the majority of applications of the BK inequality, we shall use the following version, obtained immediately from Theorem (2.12) by specializing to the case of the probability space $(\Omega, \mathcal{F}, P_p)$ of bond percolation on $\mathbb{L}^d$.

(2.15) Theorem. *Consider bond percolation on $\mathbb{L}^d$ and let A and B be increasing events defined in terms of the states of only finitely many edges. Then*

$$P_p(A \circ B) \leq P_p(A)P_p(B). \tag{2.16}$$

The condition that the events depend on only finitely many edges is easily dropped in most potential applications, and we describe how to do this in a particular important context. Let $\Pi_1, \Pi_2, \ldots, \Pi_k$ be collections of paths in $\mathbb{L}^d$; remember that paths have finite length. Let $A_i(n)$ be the event that some path in Π_i, entirely contained in the box $B(n)$, is open; we write $A_i = \lim_{n\to\infty} A_i(n)$ for the event that some path in Π_i is open. It is not difficult to see that, as $n \to \infty$,

$$A_1(n) \circ A_2(n) \circ \cdots \circ A_k(n) \to A_1 \circ A_2 \circ \cdots \circ A_k.$$

We apply (2.14) and let $n \to \infty$ to deduce that

$$P_p\big(\text{there exist \underline{disjoint} open paths } \pi_1 \in \Pi_1, \ldots, \pi_k \in \Pi_k\big) \leq P_p(A_1) \ldots P_p(A_k), \tag{2.17}$$

in contrast to the conclusion of the FKG inequality that

$$P_p\big(\text{there exist open paths } \pi_1 \in \Pi_1, \ldots, \pi_k \in \Pi_k\big) \geq P_p(A_1) \ldots P_p(A_k). \tag{2.18}$$

Van den Berg and Kesten conjectured further that a version of the BK inequality is valid for all pairs A, B of events, regardless of whether or not they are increasing. This conjecture was proved by Reimer (1997). We perform a preliminary skirmish in advance of stating Reimer's inequality formally. Let A and B be events of the probability space $(\Gamma, \mathcal{G}, P)$. For $\omega = (\omega_1, \omega_2, \ldots, \omega_m) \in \Gamma$ and $K \subseteq \{1, 2, \ldots, m\}$ we define the cylinder event $C(\omega, K)$ generated by ω on K by

$$C(\omega, K) = \{\omega' \in \Gamma : \omega'_i = \omega_i \text{ for } i \in K\}.$$

We now define $A \,\square\, B$ to be the set of all vectors $\omega \in \Gamma$ for which there exists a set $K \subseteq \{1, 2, \ldots, m\}$ such that $C(\omega, K) \subseteq A$ and $C(\omega, \overline{K}) \subseteq B$, where $\overline{K} = \{1, 2, \ldots, m\} \setminus K$.

(2.19) Theorem. Reimer's inequality. *If* $A, B \in \mathcal{F}$, *then*

$$P(A \,\square\, B) \le P(A)P(B).$$

It is immediate that $A \,\square\, B = A \circ B$ if A and B are increasing events, and a little thought leads to the observation that $A \,\square\, B = A \cap B$ if A is increasing and B is decreasing. By applying Reimer's inequality to the events A and $\overline{B}$, where A and B are increasing, we obtain that $P(A \cap B) \ge P(A)P(B)$. Therefore, Reimer's inequality includes both the FKG and BK inequalities. The hunt is on for applications of Reimer's inequality which cannot be achieved without it. See Reimer (1997) for a proof of Theorem (2.19).

There are various proofs of the BK inequality, of which we shall present possibly the easiest. We motivate this proof by describing the intuitive approach of van den Berg (1985). Let G be a finite graph and let $0 \le p \le 1$. We declare each edge of G to be open with probability p and closed otherwise, independently of all other edges. Let u, v, x, y be vertices of G, and let A and B be the events that there exist open paths from u to v and from x to y, respectively. We concentrate on the event $A \circ B$ that there exist disjoint open paths joining u to v and joining x to y. Let e be an edge of G. We replace e by two parallel edges e' and e'', each of which is open with probability p, independently of all other edges. Suppose that, in this new graph, we look for open paths from u to v which do not use e'', and we look for open paths from x to y which do not use e'. Some thought suggests that this 'splitting' of e cannot decrease the probability of disjoint open paths from u to v and from x to y. We now 'split' each edge of G in turn, replacing each edge f by two parallel edges f' and f''. At each stage we seek two disjoint paths, the first being from u to v not using edges marked $''$, and the second being from x to y not using edges marked $'$. The probability of two such disjoint paths either increases or remains unchanged at each stage. When all the edges of G have been split, we are left with two independent copies of G, in the first of which we look for an open path from u to v and in the second of which we look for an open path from x to y; such paths occur independently, so that the probability that they both occur is $P_p(A)P_p(B)$. Thus $P_p(A \circ B) \le P_p(A)P_p(B)$ as required. We now write this out formally.

Proof of BK inequality. We first produce two copies of the probability space $(\Gamma, \mathcal{F}, P)$, and we denote these copies by $(\Gamma_1, \mathcal{F}_1, P_1)$ and $(\Gamma_2, \mathcal{F}_2, P_2)$. We are concerned with the product space $(\Gamma_1 \times \Gamma_2, \mathcal{F}_2 \times \mathcal{F}_2, P_{12})$ where $P_{12} = P_1 \times P_2$, and we write $x \times y$ for a typical point in $\Gamma_1 \times \Gamma_2$; thus $x = (x_1, x_2, \dots, x_m)$ and $y = (y_1, y_2, \dots, y_m)$ where each x_i and y_j equals either 0 or 1.

Let A and B be increasing events in $\mathcal{F}$. We define the following events in the product space: we write A' for the set of all points $x \times y \in \Gamma_1 \times \Gamma_2$ for which $x \in A$, and, for $0 \le k \le m$, we write B'_k for the set of points $x \times y$ for which the composite vector $(y_1, y_2, \dots, y_k, x_{k+1}, \dots, x_m)$ belongs to B. We note that A' and B'_k are increasing events in the product space. For each point $x \times y \in \Gamma_1 \times \Gamma_2$, we say that the subset I of $\{1, 2, \dots, m\}$ *forces* A' if $u \times v \in A'$

whenever $u \times v \in \Gamma_1 \times \Gamma_2$ and $u_i = x_i$ for each $i \in I$; similarly, we say that I forces B'_k if $u \times v \in B'_k$ whenever $u_i = y_i$ for each $i \in I$ satisfying $i \leq k$ and $v_i = x_i$ for each $i \in I$ satisfying $i > k$.

The left side of (2.13) is

$$P(A \circ B) = P_{12}(A' \circ B'_0). \tag{2.20}$$

On the other hand A' and B'_m are defined in terms of disjoint sets of coordinates of $\Gamma_1 \times \Gamma_2$, so that $A' \circ B'_m = A' \cap B'_m$. Thus

$$\begin{aligned} P_{12}(A' \circ B'_m) &= P_{12}(A' \cap B'_m) \\ &= P_{12}(A')P_{12}(B'_m) \qquad \text{since } P_{12} \text{ is a product measure} \\ &= P(A)P(B). \end{aligned}$$

We shall therefore be finished once we have shown the inequality $P_{12}(A' \circ B'_0) \leq P_{12}(A' \circ B'_m)$, and to this end we claim that

$$P_{12}(A' \circ B'_{k-1}) \leq P_{12}(A' \circ B'_k) \qquad \text{for } 1 \leq k \leq m. \tag{2.21}$$

The remainder of the proof is devoted to proving this inequality. It is somewhat tedious and some readers may prefer to leave the matter to their intuition.

We partition $A' \circ B'_{k-1}$ into two events: $A' \circ B'_{k-1} = C_1 \cup C_2$ where

$$\begin{aligned} C_1 &= \{x \times y : A' \circ B'_{k-1} \text{ occurs regardless of the value of } x_k\}, \\ C_2 &= \{x \times y : x_k = 1 \text{ and } A' \circ B'_{k-1} \text{ occurs, but } A' \circ B'_{k-1} \text{ would not} \\ &\qquad \text{occur if the value of } x_k \text{ were changed to } 0\}. \end{aligned}$$

Thus C_2 is the event $\{A' \circ B'_{k-1}$ occurs if and only if $x_k = 1\} \cap \{x_k = 1\}$, and C_1 is the remainder of $A' \circ B'_{k-1}$. We further partition C_2 into two events: $C_2 = C'_2 \cup C''_2$ where

$$\begin{aligned} C'_2 &= C_2 \cap \{x \times y : \text{there exists } I \subseteq \{1, 2, \ldots, m\} \text{ such that } k \in I, \\ &\qquad I \text{ forces } A', \text{ and } \overline{I} \text{ forces } B'_{k-1}\}, \\ C''_2 &= C_2 \setminus C'_2, \end{aligned}$$

where $\overline{I} = \{1, 2, \ldots, m\} \setminus I$. Thus C'_2 is the sub-event of C_2 on which we may think of x_k as contributing essentially to A'. We now construct a measure-preserving injection φ which maps $A' \circ B'_{k-1}$ into $A' \circ B'_k$, and we do this as follows. For $x \times y \in \Gamma_1 \times \Gamma_2$ we define $x' \times y'$ to be the point in $\Gamma_1 \times \Gamma_2$ obtained from $x \times y$ by interchanging x_k and y_k, so that

$$\begin{aligned} x'_i &= x_i \quad \text{for } i \neq k, \quad \text{and } x'_k = y_k, \\ y'_i &= y_i \quad \text{for } i \neq k, \quad \text{and } y'_k = x_k. \end{aligned}$$

Let φ be defined on $A' \circ B'_{k-1}$ by

$$\varphi(x \times y) = \begin{cases} x \times y & \text{if } x \times y \in C_1 \cup C'_2, \\ x' \times y' & \text{if } x \times y \in C''_2. \end{cases} \tag{2.22}$$

We note first that $C_1 \subseteq A' \circ B'_k$, since if $x \times y \in C_1$ then A' and B'_{k-1} occur disjointly regardless of the values of x_k and y_k. Secondly, $C'_2 \subseteq A' \circ B'_k$, since if $x \times y \in C'_2$ then A' and B'_{k-1} occur disjointly and there exists such a disjoint occurrence in which x_k contributes essentially to A' and not therefore to B'_{k-1}; it follows from the definitions of A' and B'_k that $x \times y \in A' \circ B'_k$ also. Thirdly, if $x \times y \in C''_2$ then

$$\varphi(x \times y) = x' \times y' \in A' \circ B'_k,$$

since in this case $x_k = 1$ and there exists $I \subseteq \{1, 2, \dots, m\} \setminus \{k\}$ such that I forces A' and $\overline{I}$ forces B'_{k-1}; from the point of view of the configuration $x' \times y'$, I forces A' and $\overline{I}$ forces B'_k, giving that $x' \times y' \in A' \circ B'_k$ as required. We have shown that φ maps $A' \circ B'_{k-1}$ into $A' \circ B'_k$. Certainly, φ is one–one on $C_1 \cup C'_2$ since φ coincides with the identity map here. Also, φ is one–one on C''_2, and therefore φ fails to be an injection on $A' \circ B'_{k-1}$ only if there exists $x \times y \in C''_2$ with $\varphi(x \times y) = x' \times y' \in C_1 \cup C'_2$. Suppose that $x \times y$ is such a configuration in C''_2. We cannot have $x' \times y' \in C_1$, since this would imply that the values of x_k and y_k are irrelevant to the occurrence of $A' \circ B'_{k-1}$, in contradiction of the assumption that $x \times y \in C''_2$, a subset of C_2. It is impossible also that $x' \times y' \in C'_2$, since in this case $x_k = 1$ and $y_k = 1$ by virtue of the assumptions that $x \times y \in C''_2$ and $x' \times y' \in C'_2$, respectively; therefore $x' \times y' = x \times y$, which contradicts the assumption that one lies in C'_2 and the other in C''_2. It follows that φ is an injection from $A' \circ B'_{k-1}$ into $A' \circ B'_k$ as claimed.

Finally, φ preserves measure since $P_{12}(\varphi(x \times y)) = P_{12}(x \times y)$ whenever $x \times y \in A' \circ B'_{k-1}$. In conclusion, we deduce that $P_{12}(A' \circ B'_{k-1}) \leq P_{12}(A' \circ B'_k)$, and (2.21) is proved. □

2.4 Russo's Formula

We saw in Theorem (2.1) that, for increasing events A, the probability $P_p(A)$ is a non-decreasing function of p. Our third basic technique is a method for estimating the rate of change of $P_p(A)$ as a function of p. We may motivate the argument as follows. First, in comparing $P_p(A)$ and $P_{p+\delta}(A)$, it is useful to construct the two processes having densities p and $p + \delta$ on the same probability space in the usual way. Let $(X(e) : e \in \mathbb{E}^d)$ be independent random variables having the uniform distribution on $[0, 1]$, and define $\eta_p(e) = 1$ if $X(e) < p$ and $\eta_p(e) = 0$ otherwise. For increasing events A,

$$P_{p+\delta}(A) - P_p(A) = P\big(\eta_p \notin A,\ \eta_{p+\delta} \in A\big). \tag{2.23}$$

If $\eta_p \notin A$ but $\eta_{p+\delta} \in A$, there must exist edges e satisfying $\eta_p(e) = 0$, $\eta_{p+\delta}(e) = 1$, which is to say that $p \leq X(e) < p + \delta$. Let $E_{p,\delta}$ be the set of such edges, and assume that A depends on the states of only finitely many edges. Clearly $P(|E_{p,\delta}| \geq 2) = o(\delta)$ as $\delta \downarrow 0$, and so we shall neglect the possibility that there exist two or more edges in $E_{p,\delta}$. If e is the unique edge satisfying $p \leq X(e) < p + \delta$, then e must be 'essential' for A in the sense that $\eta_p \notin A$ but $\eta'_p \in A$ where η'_p is obtained from η_p by changing the state of e from 0 to 1. The 'essentialness' of e does not depend on the state of e, so that each edge e contributes roughly an amount

(2.24)
$$P\big(p \leq X(e) < p + \delta\big) P_p(e \text{ is 'essential' for } A) = \delta P_p(e \text{ is 'essential' for } A)$$

towards the quantity in (2.23). We divide by δ and take the limit as $\delta \downarrow 0$ to obtain without rigour the formula

$$\frac{d}{dp} P_p(A) = \sum_e P_p(e \text{ is 'essential' for } A).$$

We shall now derive a rigorous version of this formula, valid for all increasing events A which depend on only finitely many edges.

Let A be an event, not necessarily increasing, and let ω be a configuration. We call the edge e *pivotal* for the pair (A, ω) if $I_A(\omega) \neq I_A(\omega')$, where ω' is the configuration which agrees with ω on all edges except e, and $\omega'(e) = 1 - \omega(e)$. Thus e is pivotal for (A, ω) if the occurrence or non-occurrence of A depends crucially on the state of e. The event that 'e is pivotal for A' is the set of configurations ω for which e is pivotal for (A, ω). We note that this event depends only on the states of edges other than e; it is independent of the state of e itself. We shall be interested particularly in *increasing* events A; for such an event A, an edge e is pivotal if and only if A does not occur when e is closed but A does occur when e is open.

Here are two examples:

(a) Let A be the event that the origin lies in an infinite open cluster. The edge e is pivotal for A if, when e is removed from the lattice, one endvertex of e is in a finite open cluster which includes the origin, and the other endvertex of e is in an infinite open cluster.

(b) Let A be the event that there exists an open path within the box $B(n)$ in two dimensions, joining some vertex on the left side of $B(n)$ to some vertex on the right side (see Figure 2.1). The edge e of $B(n)$ is pivotal for A if, when e is removed from the lattice, no such open crossing of $B(n)$ exists, but one endvertex of e is joined by an open path to the left side of $B(n)$ and the other endvertex of e is joined similarly to the right side of $B(n)$.

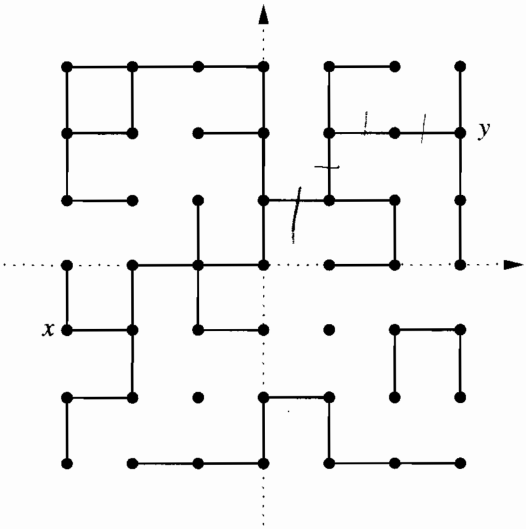


Figure 2.1. The box $B(3)$, with an open path joining a vertex x on the left side to a vertex y on the right side. This configuration has four edges which are pivotal for the existence of an open path joining the left side to the right side.

(2.25) Theorem. Russo's formula. *Let A be an increasing event defined in terms of the states of only finitely many edges of $\mathbb{L}^d$. Then*

$$\frac{d}{dp} P_p(A) = E_p\big(N(A)\big), \tag{2.26}$$

where $N(A)$ is the number of edges which are pivotal for A.

Formula (2.26) may be written as

$$\frac{d}{dp} P_p(A) = \sum_{e \in \mathbb{E}^d} P_p(e \text{ is pivotal for } A). \tag{2.27}$$

Such formulae are not generally valid for events which depend on more than finitely many edges; indeed $P_p(A)$ need not be differentiable for all values of p (for example, we shall remark in the notes for Section 10.2 that, for percolation in two dimensions, $(\theta(p) - \theta(p_c))/(p - p_c)$ equals 0 if $p < p_c$ but tends to ∞ as $p \downarrow p_c$). For general increasing events A, the best that the method allows is a lower bound on the right-hand derivative of $P_p(A)$:

$$\liminf_{\delta \downarrow 0} \frac{1}{\delta} \big(P_{p+\delta}(A) - P_p(A)\big) \geq E_p\big(N(A)\big). \tag{2.28}$$

The event $\{e \text{ is pivotal for } A\}$ is independent of the state of e, and thus

$$P_p(e \text{ is pivotal for } A) = \frac{1}{p} P_p(e \text{ is open and pivotal for } A).$$

Hence, if A is increasing and depends on only finitely many edges,

$$\begin{aligned}
(2.29)\qquad \frac{d}{dp} P_p(A) &= \frac{1}{p} \sum_e P_p(e \text{ is open and pivotal for } A) \\
&= \frac{1}{p} \sum_e P_p\big(A \cap \{e \text{ is pivotal for } A\}\big) \\
&= \frac{1}{p} \sum_e P_p\big(e \text{ is pivotal for } A \bigm| A\big) P_p(A) \\
&= \frac{1}{p} E_p\big(N(A) \bigm| A\big) P_p(A).
\end{aligned}$$

We may make use of this expression in the following way. Divide throughout by $P_p(A)$ and integrate over the interval $[p_1, p_2]$ to obtain
(2.30)

$$P_{p_2}(A) = P_{p_1}(A) \exp\left(\int_{p_1}^{p_2} \frac{1}{p} E_p\big(N(A) \bigm| A\big)\, dp\right) \quad \text{if } 0 \le p_1 < p_2 \le 1.$$

Applications of this expression require estimates for $E_p(N(A) \mid A)$. Accurate estimates for the probability that a given edge e is pivotal are often rather hard to derive, and it is more common that we resort to an inequality of the FKG or BK type in this context. If we use the trivial bound

$$P_p\big(e \text{ is pivotal for } A \bigm| A\big) \le 1,$$

then (2.30) yields

$$(2.31)\qquad P_{p_2}(A) \le (p_2/p_1)^m P_{p_1}(A) \qquad \text{if } 0 \le p_1 < p_2 \le 1,$$

where m is the (finite) number of edges in terms of which A is defined. This inequality is a quantified version of the statement that $P_p(A)$ cannot increase 'too fast' as p increases. We shall make use of it later.

Proof of Russo's formula and Equation (2.28). We shall make rigorous the discussion at the beginning of the section. Let A be an increasing event and let $\mathbf{p} = (p(e) : e \in \mathbb{E}^d)$ be a collection of numbers satisfying $0 \le p(e) \le 1$ for all e. Let $(X(e) : e \in \mathbb{E}^d)$ be independent random variables each having the uniform distribution on $[0, 1]$. We construct the configuration $\eta_{\mathbf{p}}$ on $\mathbb{E}^d$ by defining $\eta_{\mathbf{p}}(e) = 1$ if $X(e) < p(e)$ and $\eta_{\mathbf{p}}(e) = 0$ otherwise. Writing $P_{\mathbf{p}}$ for the probability measure on Ω in which the state $\omega(e)$ of the edge e equals 1 with probability $p(e)$, we have that

$$P_{\mathbf{p}}(A) = P(\eta_{\mathbf{p}} \in A)$$

as usual. Now, choose an edge f and define $\mathbf{p}' = (p'(e) : e \in \mathbb{E}^d)$ by

$$p'(e) = \begin{cases} p(e) & \text{if } e \ne f, \\ p'(f) & \text{if } e = f; \end{cases}$$

$\mathbf{p}$ and $\mathbf{p}'$ may differ only at the edge f. Now, if $p(f) \le p'(f)$ then

$$\begin{aligned} P_{\mathbf{p}'}(A) - P_{\mathbf{p}}(A) &= P(\eta_{\mathbf{p}} \notin A,\ \eta_{\mathbf{p}'} \in A) \\ &= \big\{p'(f) - p(f)\big\} P_{\mathbf{p}}(f \text{ is pivotal for } A), \end{aligned}$$

by the discussion leading to (2.24). We divide by $p'(f) - p(f)$ and take the limit as $p'(f) - p(f) \to 0$ to obtain

$$\frac{\partial}{\partial p(f)} P_{\mathbf{p}}(A) = P_{\mathbf{p}}(f \text{ is pivotal for } A).$$

So far we have assumed nothing about A save that it be increasing. If A depends on finitely many edges only, then $P_{\mathbf{p}}(A)$ is a function of a finite collection $(p(f_i) : 1 \le i \le m)$ of edge-probabilities, and the chain rule gives that

$$\begin{aligned} \frac{d}{dp} P_p(A) &= \sum_{i=1}^{m} \frac{\partial}{\partial p(f_i)} P_{\mathbf{p}}(A) \bigg|_{\mathbf{p}=(p,p,\dots,p)} \\ &= \sum_{i=1}^{m} P_p(f_i \text{ is pivotal for } A) \\ &= E_p(N(A)) \end{aligned}$$

as required. If, on the other hand, A depends on infinitely many edges, we let E be a finite collection of edges and define

$$\mathbf{p}_E(e) = \begin{cases} p & \text{if } e \notin E, \\ p + \delta & \text{if } e \in E, \end{cases}$$

where $0 \le p \le p + \delta \le 1$. Now, A is increasing, so that

$$P_{p+\delta}(A) \ge P_{\mathbf{p}_E}(A),$$

and hence

$$\begin{aligned} \frac{1}{\delta}\big(P_{p+\delta}(A) - P_p(A)\big) &\ge \frac{1}{\delta}\big(P_{\mathbf{p}_E}(A) - P_p(A)\big) \\ &\to \sum_{e \in E} P_p(e \text{ is pivotal for } A) \end{aligned}$$

as $\delta \downarrow 0$. We let $E \uparrow \mathbb{E}^d$ to obtain (2.28). □

The above argument may be extended in various ways. We begin with some notation. For $\omega \in \Omega$ and $A, B \subseteq \mathbb{E}^d$ with $A \cap B = \varnothing$, let ω_B^A be the configuration given by

$$\omega_B^A(e) = \begin{cases} \omega(e) & \text{if } e \notin A \cup B, \\ 1 & \text{if } e \in A, \\ 0 & \text{if } e \in B. \end{cases}$$

For ease of notation, we abbreviate singleton sets $A = \{f\}$ by f, and pairs $A = \{e, f\}$ by ef. For a random variable X, we define the increment of X at e by

$$\delta_e X(\omega) = X(\omega^e) - X(\omega_e).$$

It is an easy matter to extend Russo's formula to the following, and we omit the proof.

(2.32) Theorem. *Let X be a random variable which is defined in terms of the states of only finitely many edges of $\mathbb{L}^d$. Then*

$$\frac{d}{dp}E_p(X) = \sum_{e\in\mathbb{E}^d} E_p(\delta_e X).$$

Turning to the second derivative, it follows from the theorem that

$$\frac{d^2}{dp^2}E_p(X) = \sum_{e,f\in\mathbb{E}^d} E_p(\delta_e\delta_f X).$$

Now $\delta_e\delta_e X = 0$, and for $e \neq f$

$$\delta_e\delta_f X(\omega) = X(\omega^{ef}) - X(\omega^e_f) - X(\omega^f_e) + X(\omega_{ef}).$$

Let X be the indicator function I_A of the increasing event A. Since I_A is increasing,

(2.33)
$$\begin{aligned}\frac{d^2}{dp^2}P_p(A) = \sum_{\substack{e,f\in\mathbb{E}^d\\ e\neq f}} E_p\Big\{ & I_A(\omega^{ef})\big(1-I_A(\omega^f_e)\big)\big(1-I_A(\omega^e_f)\big) \\ & -I_A(\omega^e_f)I_A(\omega^f_e)\big(1-I_A(\omega_{ef})\big)\Big\} \\ = E_p(N^{\text{ser}}_A) - E_p(N^{\text{par}}_A) & \end{aligned}$$

where N^{ser}_A (respectively N^{par}_A) is the number of distinct ordered pairs e, f of edges such that $\omega^{ef} \in A$ but $\omega^f_e, \omega^e_f \notin A$ (respectively $\omega^f_e, \omega^e_f \in A$ but $\omega_{ef} \notin A$). (The superscripts here are abbreviations for 'series' and 'parallel'.) This argument may be generalized to higher derivatives.

2.5 Inequalities of Reliability Theory

Under the title 'reliability theory' one studies, amongst other things, the probability of a route through a network whose individual links are unreliable. More precisely, let G be a finite graph, and declare each edge of G to be open with probability p, independently of all other edges. Let x and y be specified vertices of G and let $h(p)$ be the probability that there exists an open path of G from x to y. In studying the behaviour of h as a function of p, reliability theorists have developed various tools which are relevant to percolation. These include versions of the FKG inequality and Russo's formula; see Moore and Shannon (1956), Esary and Proschan (1963), and Barlow and Proschan (1965, 1975). We present some of their results here, in forms suitable for application to percolation.

Let E be a finite set of edges of $\mathbb{L}^d$. We shall confine ourselves to events which depend only on the edges in E. We write $\text{cov}_p(X, Y)$ for the covariance of two random variables X and Y under the measure P_p.

(2.34) Theorem. *Let A be an event which depends only on the edges in E, and let N be the (random) number of edges of E which are open. Then*

$$\frac{d}{dp} P_p(A) = \frac{1}{p(1-p)} \operatorname{cov}_p(N, I_A) \qquad \text{for } 0 < p < 1. \tag{2.35}$$

This theorem is valid for all events A regardless of whether or not they are monotone. The following is an immediate corollary, of which the second part is the main step in the celebrated 'S-shape' theorem of reliability theory.

(2.36) Theorem. *Let A be an event which depends only on the edges in E, and suppose that $0 < p < 1$. Then*

(a) $$\frac{d}{dp} P_p(A) \le \sqrt{\frac{m P_p(A)(1 - P_p(A))}{p(1-p)}} \qquad \text{where } m = |E|,$$

(b) *if A is increasing, we have that*

$$\frac{d}{dp} P_p(A) \ge \frac{P_p(A)(1 - P_p(A))}{p(1-p)}. \tag{2.37}$$

Before proving these results, we mention a related inequality which we shall use later.

(2.38) Theorem. *Let A be an increasing event which depends on only finitely many edges of $\mathbb{L}^d$, and suppose that $0 < p < 1$. Then $\log P_p(A)/\log p$ is a non-increasing function of p.*

This last theorem amounts to saying that $h(p) = P_p(A)$ satisfies

$$h(p^\gamma) \le h(p)^\gamma \qquad \text{if } 0 < p < 1 \text{ and } \gamma \ge 1, \tag{2.39}$$

whenever A is increasing. Rewriting the conclusion of the theorem in terms of the derivative of $\log P_p(A)/\log p$, we obtain

$$\frac{d}{dp} P_p(A) \ge \frac{P_p(A) \log P_p(A)}{p \log p}. \tag{2.40}$$

This is an improvement over inequality (2.37) whenever $P_p(A) < p$.

Proof of Theorem (2.34). We write ω for a configuration of the edges in E, and $N(\omega)$ for the number of open edges of ω. Clearly

$$P_p(A) = \sum_{\omega} I_A(\omega) p^{N(\omega)} q^{m - N(\omega)},$$

where $p + q = 1$ and $m = |E|$. Thus

$$\begin{aligned}\frac{d}{dp} P_p(A) &= \sum_\omega I_A(\omega) p^{N(\omega)} q^{m-N(\omega)} \left(\frac{N(\omega)}{p} - \frac{m - N(\omega)}{q} \right) \\ &= \frac{1}{pq} \sum_\omega I_A(\omega) p^{N(\omega)} q^{m-N(\omega)} \big(N(\omega) - mp\big) \\ &= \frac{1}{pq} \operatorname{cov}_p(N, I_A)\end{aligned}$$

since $E_p(N) = mp$. □

Proof of Theorem (2.36). We apply the Cauchy–Schwarz inequality to the right side of (2.35) to obtain

$$\frac{d}{dp} P_p(A) \le \frac{1}{p(1-p)} \sqrt{\operatorname{var}_p(N) \operatorname{var}_p(I_A)}$$

and part (a) follows from the observation that N has the binomial distribution with parameters m and p, and I_A has a Bernoulli distribution.

Suppose now that A is increasing. Then

(2.41) $$\begin{aligned}\operatorname{cov}_p(N, I_A) &= \operatorname{cov}_p(I_A, I_A) + \operatorname{cov}_p(N - I_A, I_A) \\ &\ge \operatorname{var}_p(I_A)\end{aligned}$$

by the FKG inequality, since $N - I_A$ is an increasing random variable. We substitute this into (2.35) to obtain (2.37). □

Proof of Theorem (2.38). We shall show that $h(p) = P_p(A)$ satisfies

(2.42) $$h(p^\gamma) \le h(p)^\gamma \qquad \text{if } 0 < p < 1 \text{ and } \gamma \ge 1,$$

whenever A is increasing and depends on only finitely many edges of $\mathbb{L}^d$. To see that this implies the result, substitute $p_1 = p^\gamma$ in (2.42) and take logarithms to find that

$$\frac{\log h(p_1)}{\log p_1} \ge \frac{\log h(p)}{\log p}$$

as required.

We shall prove (2.42) by induction on the number of edges which are relevant to A. Suppose that A is increasing and depends only on the edges in the finite set E, and write $m = |E|$. If $m = 1$ then A is either $\{0, 1\}$ or $\{1\}$, and so $h(p) = P_p(A)$ equals either 1 or p; in either case (2.42) is valid. Suppose therefore that k (≥ 1) is such that (2.42) is valid whenever $m \le k$, and consider the case when $m = k + 1$. Let $0 < p < 1$, $\gamma \ge 1$, and let e be an edge in E. Then

(2.43) $$\begin{aligned}h(p^\gamma) &= P_{p^\gamma}\big(A \,\big|\, \omega(e) = 1\big) p^\gamma + P_{p^\gamma}\big(A \,\big|\, \omega(e) = 0\big)(1 - p^\gamma) \\ &\le P_p\big(A \,\big|\, \omega(e) = 1\big)^\gamma p^\gamma + P_p\big(A \,\big|\, \omega(e) = 0\big)^\gamma (1 - p^\gamma)\end{aligned}$$

by the induction hypothesis. It is not difficult to check that

$$x^\gamma p^\gamma + y^\gamma(1-p^\gamma) \le \{xp + y(1-p)\}^\gamma \qquad \text{when } x \ge y \ge 0; \tag{2.44}$$

to see this, check that equality holds when $x = y \ge 0$ and that the derivative of the left side with respect to x is at most the corresponding derivative of the right side when $x, y \ge 0$. We note that $P_p(A \mid \omega(e) = 1) \ge P_p(A \mid \omega(e) = 0)$ since A is increasing, and (2.43) now yields

$$\begin{aligned} h(p^\gamma) &\le \Big\{P_p\big(A \,\big|\, \omega(e) = 1\big)p + P_p\big(A \,\big|\, \omega(e) = 0\big)(1-p)\Big\}^\gamma \\ &= h(p)^\gamma \end{aligned}$$

as required. □

2.6 Another Inequality

There is a case for delaying the presentation of our final inequality until it is needed, but we include it here with the recommendation that the reader note its existence and move on.

Let r be a positive integer. For any configuration ω of edge-states, we define the *sphere* with radius r and centre at ω by

$$S_r(\omega) = \left\{\omega' \in \Omega : \sum_{e \in \mathbb{E}^d} |\omega'(e) - \omega(e)| \le r\right\};$$

$S_r(\omega)$ is the collection of configurations which differ from ω on at most r edges. For any event A, we define the *interior* of A with depth r by

$$I_r(A) = \{\omega \in \Omega : S_r(\omega) \subseteq A\};$$

thus $I_r(A)$ is the set of configurations in A which are still in A even if we perturb the states of up to r edges.

(2.45) Theorem. *Let A be an increasing event and let r be a positive integer. Then*

$$1 - P_{p_2}\big(I_r(A)\big) \le \left(\frac{p_2}{p_2 - p_1}\right)^r \big\{1 - P_{p_1}(A)\big\} \tag{2.46}$$

whenever $0 \le p_1 < p_2 \le 1$.

We may think of $I_r(A)$ as the event that A occurs and is 'stable' with respect to changes in the states of r or fewer edges. The theorem amounts to the assertion that, if A is likely to occur when the edge-probability is p_1, then $I_r(A)$ is likely to occur when the edge-probability exceeds p_1. Such an inequality is valuable in the context of the following type of construction. Let A_n be the event that there exists an open path between the left and right sides of the rectangle $[0, n]^2$ in two dimensions; such a path is called a 'left–right crossing'. In this case $I_r(A_n)$ is the event that such a path exists irrespective of the states of any collection of r edges. Those familiar with the max-flow min-cut theorem of network flows (see Section 13.1) and Menger's theorem (see Wilson (1979, p. 126)) will recognize $I_r(A_n)$ as the event that there exist at least $r + 1$ edge-disjoint left–right crossings of $[0, n]^2$. We shall see at the end of Section 11.3 that $P_p(A_n) \geq 1 - e^{-n\alpha(p)}$ for some $\alpha(p) > 0$, so long as $p > \frac{1}{2}$. We apply (2.46) with $r = \beta n$ to find that

$$(2.47) \qquad P_{p_2}(K_n \leq \beta n) \leq \left(\frac{p_2}{p_2 - p_1}\right)^{\beta n} e^{-n\alpha(p_1)},$$

where K_n is the maximal number of edge-disjoint left–right crossings of the rectangle, and $\frac{1}{2} < p_1 < p_2 \leq 1$. If

$$\gamma(p_1, p_2, \beta) = \beta \log\left(\frac{p_2}{p_2 - p_1}\right) - \alpha(p_1)$$

satisfies $\gamma(p_1, p_2, \beta) < 0$, the left side of (2.47) decays exponentially as $n \to \infty$. We have proved that, if for some value of p there is an exponentially decaying probability that there exists no left–right crossing of $[0, n]^2$, then for larger values of p there is an exponentially decaying probability that there are $o(n)$ such crossings.

Inequality (2.46) is one of several related inequalities, of which another follows. Let A be an increasing event. For $\omega \in \Omega$, let $F_A(\omega)$ denote the 'distance' of ω from A, that is,

$$(2.48) \qquad F_A(\omega) = \inf\left\{\sum_e \big(\omega'(e) - \omega(e)\big) : \omega' \geq \omega,\ \omega' \in A\right\}.$$

Note that $F_A(\omega) = 0$ if $\omega \in A$. One may now prove in much the same way as (2.46) that

$$(2.49) \qquad P_{p_2}(A) \geq \left(\frac{p_2 - p_1}{1 - p_1}\right)^r P_{p_1}(F_A \leq r) \quad \text{if } p_1 \leq p_2 \text{ and } r \geq 0.$$

Proof of Theorem (2.45). We follow Aizenman, Chayes, Chayes, Fröhlich, and Russo (1983). Let $(X(e) : e \in \mathbb{E}^d)$ and $(\eta_p : 0 \leq p \leq 1)$ be defined in the usual way; see Section 1.3. Suppose that $0 \leq p_1 < p_2 \leq 1$ and that A is an increasing event. If $\eta_{p_2} \notin I_r(A)$, there exists a (random) collection $B = B(\eta_{p_2})$ of edges such that:

(a) $|B| \le r$;

(b) $\eta_{p_2}(e) = 1$ for all $e \in B$; and

(c) the configuration η, obtained from η_{p_2} by declaring all edges in B to be closed, satisfies $\eta \notin A$.

There may exist many such sets B, in which case we choose the earliest in some fixed ordering of all possible such sets. Suppose now that $\eta_{p_2} \notin I_r(A)$, and that every edge e in the set B satisfies $p_1 \le X(e) < p_2$; it follows from (c) above that $\eta_{p_1} \notin A$. Conditional on B, there is probability $\{(p_2 - p_1)/p_2\}^{|B|}$ that $p_1 \le X(e) < p_2$ for all $e \in B$, and therefore

$$P\big(\eta_{p_1} \notin A \,\big|\, \eta_{p_2} \notin I_r(A)\big) \ge \left(\frac{p_2 - p_1}{p_2}\right)^r$$

since $|B| \le r$. Inequality (2.46) follows easily. □

2.7 Notes

Section 2.2. The second inequality of Theorem (2.4) might well be dubbed Harris's inequality (or the Harris–FKG inequality) after Harris (1960); a version of this inequality is standard in reliability theory, and may be found in Barlow and Proschan (1965, p. 207) for example. There is a more general version due to Fortuin, Kasteleyn, and Ginibre (1971) which is of great value in statistical physics. For a general mathematical approach to such correlation inequalities, see Ahlswede and Daykin (1979), Batty and Bollman (1980), den Hollander and Keane (1986), and the references therein. The proof given here is derived from material in Kesten (1982) and Durrett (1988, Chapter 6).

Section 2.3. The BK inequality was first stated and proved in a more general form by van den Berg and Kesten (1985). There are several refinements in van den Berg and Fiebig (1987). The proof given here has its origins in the work of McDiarmid (1980, 1981, 1983). Van den Berg (1985) has pointed out that a similar proof may be used to show the FKG inequality. Similar ideas occur in papers of Hammersley (1961), Rüschendorf (1982), and Campanino and Russo (1985). Reimer's inequality may be found in Reimer (1997).

Section 2.4. Russo's formula was proved by Russo (1981). Such techniques are well known in reliability theory. For example, Russo's formula is essentially equation (4.4) of Barlow and Proschan (1965, p. 210). There are several interesting approaches to its proof; see Chayes and Chayes (1986a) for example. Equation (3.15) of Kesten (1981) is a special case of inequality (2.31).

Section 2.5. Theorem (2.34) may be found in Barlow and Proschan (1965, p. 210). The inequalities of Theorem (2.36) appear in Moore and Shannon (1956). Part

(a) has been rediscovered by Chayes, Chayes, Fisher, and Spencer (1986), and our proof resembles theirs. The claim of Theorem (2.38) was suggested by J. van den Berg.

Section 2.6. The inequality of Theorem (2.45) was proved by Aizenman, Chayes, Chayes, Fröhlich, and Russo (1983), and has been put to use by Chayes and Chayes (1986a, d). It refers to the effect of adding a low density of extra open edges. This technique is known to some as 'sprinkling', and is of special value when studying the supercritical phase of percolation; see Section 7.2. In the more general context of random-cluster models, Grimmett (1997) has proved inequality (2.49) together with the complementary differential inequality

$$\frac{d}{dp}\{\log P_p(A)\} \geq \frac{E_p(F_A)}{p(1-p)},$$

valid for non-empty, increasing, cylinder events A. These inequalities have applications to entanglement in percolation; see Section 12.5.

The 'concentration inequalities' of Talagrand and others have found applications in areas very closely related to percolation. Such inequalities are not included here, and the reader is referred to Talagrand (1995) for a recent account. Similarly, the 'threshold theorems' of Friedgut and Kalai (1996) and others may be applied to percolation events on finite grids.

Chapter 3

Critical Probabilities

3.1 Equalities and Inequalities

Let G be a graph, and let $p_c(G)$ denote the critical probability of bond percolation on G, as in Section 1.6. It is tempting to seek an exact calculation of $p_c(G)$ for given G, but there seems no reason to expect a closed form for $p_c(G)$ unless G has special structure. Indeed, except for certain famous two-dimensional lattices, the value of $p_c(G)$ may have no special numerical features. The exceptional cases include:

square lattice	$p_c = \frac{1}{2}$
triangular lattice	$p_c = 2\sin(\pi/18)$
hexagonal lattice	$p_c = 1 - 2\sin(\pi/18)$
bow-tie lattice	$p_c = p_c(\text{bow-tie})$

where $p_c(\text{bow-tie})$ is the unique root in $(0, 1)$ of the equation

$$1 - p - 6p^2 + 6p^3 - p^5 = 0.$$

See Figure 3.1 for drawings of these lattices.

It is the operation of 'duality' which is of primary value in establishing these exact values (the definition of planar dual is given beneath (1.16), see also Section 11.2). Given a planar lattice $\mathcal{L}$ (defined in an appropriate way not explored here) and its dual lattice $\mathcal{L}_d$, one may show that

$$p_c(\mathcal{L}) + p_c(\mathcal{L}_d) = 1 \tag{3.1}$$

subject to certain conditions of symmetry on $\mathcal{L}$. We do not present a proof of such a relation, since this would use techniques to be explored only later in this book. Equation (3.1) is equivalent to the following statement:

$$p > p_c(\mathcal{L}) \quad \text{if and only if} \quad 1 - p < p_c(\mathcal{L}_d),$$

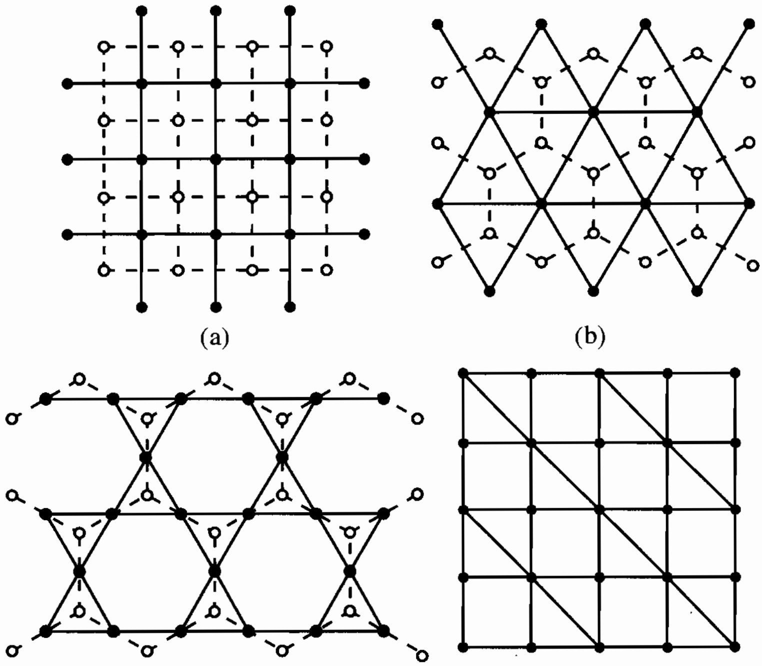


Figure 3.1. Five lattices in two dimensions. (a) The square lattice is self-dual. (b) The broken edges constitute the hexagonal (or 'honeycomb') lattice, and the solid edges constitute the triangular lattice, being the dual of the hexagonal lattice. (c) The broken edges constitute the hexagonal lattice, and the solid edges constitute the 'kagomé' lattice, being the covering lattice of the hexagonal lattice. (d) The so called bow-tie lattice.

for which an intuitive explanation is as follows. If $p > p_c(\mathcal{L})$, there exists (almost surely) an infinite open cluster of $\mathcal{L}$, and infinite clusters occupy a strictly positive density of space. If there is a unique such infinite cluster (which fact we shall prove in Chapter 8), then this cluster extends throughout space, and precludes the existence of an infinite closed cluster of $\mathcal{L}_d$; therefore $1 - p < p_c(\mathcal{L}_d)$. Conversely, if $p < p_c(\mathcal{L})$, all open clusters of $\mathcal{L}$ are (almost surely) finite, and the intervening space should contain an infinite closed cluster of $\mathcal{L}_d$; therefore, $1-p > p_c(\mathcal{L}_d)$. However appealing these crude arguments may be, their rigorous justification is highly non-trivial.

Once (3.1) is accepted, the exact value $p_c = \frac{1}{2}$ for the square lattice follows immediately, since this lattice is self-dual. A rigorous proof of this calculation appears in Section 11.3. When $\mathcal{L}$ is the triangular lattice, then $\mathcal{L}_d$ is the hexagonal lattice, and in this case we need another link between the two critical probabilities in order to compute them exactly. The so called 'star–triangle' relation provides such a link, and the exact values follow. See Section 11.9 for a complete derivation.

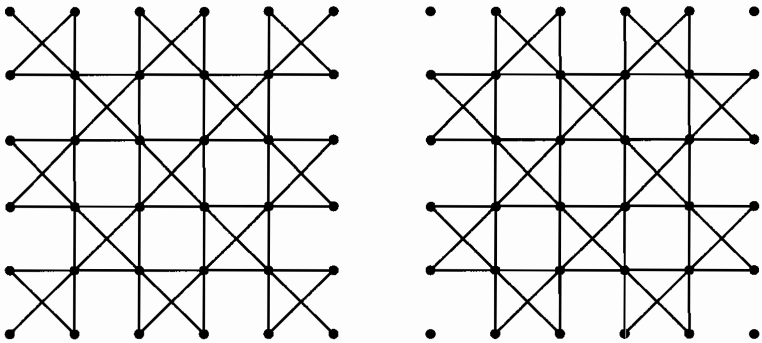


Figure 3.2. A graph G may be used to generate a matching pair $\mathcal{G}_1$, $\mathcal{G}_2$. Any finite cluster of $\mathcal{G}_1$ is surrounded by a circuit of $\mathcal{G}_2$. In this picture, G is the square lattice $\mathbb{L}^2$, and $\mathcal{G}_1$ and $\mathcal{G}_2$ are obtained by adding the diagonals to alternate faces of G. In this special case, both $\mathcal{G}_1$ and $\mathcal{G}_2$ are isomorphic to the covering lattice of $\mathbb{L}^2$.

A similar argument is valid for the bow-tie lattice $\mathcal{L}$, namely that the dual of $\mathcal{L}$ may be transformed into a copy of $\mathcal{L}$, by judicious use of the star–triangle transformation. This enables a computation of its critical value. There may exist other two-dimensional lattices to which similar arguments may be applied.

We turn now to site percolation. As observed in Section 1.6, the bond model on a graph G is equivalent to the site model on the covering graph G_c. It follows in particular that the kagomé lattice, being the covering lattice of the hexagonal lattice, satisfies $p_c^{\text{site}}(\text{kagomé}) = 1 - 2\sin(\pi/18)$.

Whereas *duality* was a key to *bond* percolation in two dimensions, the corresponding property for *site* percolation is that of *matching*. A matching pair $\mathcal{G}_1$, $\mathcal{G}_2$ of graphs in two dimensions is constructed as follows. We begin with an infinite planar graph G with 'origin' 0, and we select some arbitrary family $\mathcal{F}$ of faces of G. We obtain $\mathcal{G}_1$ (respectively $\mathcal{G}_2$) from G by adding all diagonals to all faces in $\mathcal{F}$ (respectively all faces not in $\mathcal{F}$). The graphs G, $\mathcal{G}_1$, $\mathcal{G}_2$ have the same vertex sets, and therefore a site percolation process on G induces site percolation processes on $\mathcal{G}_1$ and $\mathcal{G}_2$. If the origin 0 belongs to a finite open cluster of $\mathcal{G}_1$, then the external (vertex) boundary of this cluster forms a closed circuit of $\mathcal{G}_2$ (see the example in Figure 3.2). This turns out to be a very useful property. We say that $\mathcal{G}_1$ is *self-matching* if $\mathcal{G}_1$ and $\mathcal{G}_2$ are isomorphic graphs. Note that, if G is a triangulation (that is, if every face of G is a triangle), then $G = \mathcal{G}_1 = \mathcal{G}_2$, and in this case G is self-matching. The triangular lattice $\mathbb{T}$ is an example of a self-matching lattice. Further details and references concerning two-dimensional matching pairs may be found in Kesten (1982).

Let $\mathcal{G}_1$, $\mathcal{G}_2$ be a matching pair of lattices in two dimensions. Subject to assumptions on the pair $\mathcal{G}_1$, $\mathcal{G}_2$, one may on occasion be able to justify the relation

$$p_c^{\text{site}}(\mathcal{G}_1) + p_c^{\text{site}}(\mathcal{G}_2) = 1;$$

cf. (3.1). One may deduce that the triangular lattice $\mathbb{T}$, being self-matching, has

	bond	site
hexagonal		$\simeq 0.70$
square $\mathbb{L}^2$		$\simeq 0.59$
kagomé	$\simeq 0.52$	
cubic $\mathbb{L}^3$	$\simeq 0.25$	$\simeq 0.31$

Table 3.1. Numerical estimates of critical probabilities of bond and site percolation models on four lattices. See Hughes (1996) for origins and explanations.

site critical probability $p_c^{\text{site}}(\mathbb{T}) = \frac{1}{2}$. Indeed, it is believed that $p_c^{\text{site}} = \frac{1}{2}$ for a broad family of 'reasonable' triangulations of the plane.

In the absence of a general method for computing critical percolation probabilities, we may have cause to seek inequalities. These come in two forms, rigorous and non-rigorous. A great deal of estimation of critical probabilities has been carried out, using a mixture of numerical, rigorous, and non-rigorous arguments. We do not survey such results here, but refer the reader to pages 182–183 of Hughes (1996). As an example of an inequality which is both rigorous and rather tight, Wierman (1990) has proved that

$$0.5182 \le p_c^{\text{bond}}(\text{kagomé}) \le 0.5335,$$

but other results of this type are generally rather weak.

Another line of enquiry has been to understand the behaviour of critical probabilities in the limit as the number d of dimensions is allowed to pass to infinity. We shall encounter in Section 10.3 the technology known as the 'lace expansion', which has been developed by Hara and Slade (1990, 1994, 1995) in order to understand percolation when d is large and finite. When applied to bond percolation on $\mathbb{L}^d$, these arguments imply an expansion of which the first terms follow:

$$\text{(3.2)} \qquad p_c^{\text{bond}}(\mathbb{L}^d) = \frac{1}{2d} + \frac{1}{(2d)^2} + \frac{7/2}{(2d)^3} + \mathrm{O}\left(\frac{1}{(2d)^4}\right) \quad \text{as } d \to \infty.$$

The remainder of this chapter is devoted to a method for proving strict inequalities *between* critical probabilities. This method appears to have fundamental merit in situations where one needs to understand whether a systematic addition of edges to a process causes a *strict* change in its critical value. In Section 3.2 is presented an example of this argument at work; see Theorem (3.7). Section 3.3 contains a general formulation of enhancements for percolation models. Such methods are adapted in Section 3.4 to obtain strict inequalities between site and bond critical probabilities of cubic lattices.

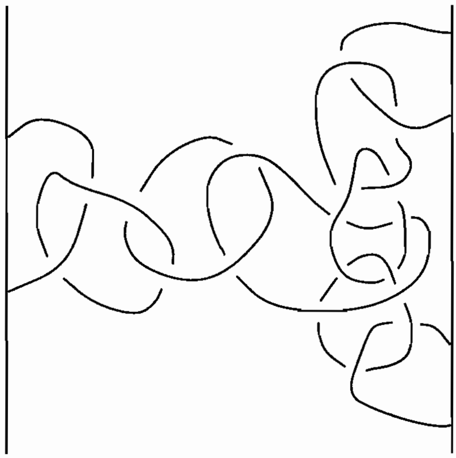


Figure 3.3. An entanglement between opposite sides of a cube in three dimensions. Note the chain of necklaces on the right.

3.2 Strict Inequalities

If $\mathcal{L}$ is a sublattice of the lattice $\mathcal{L}'$ (written $\mathcal{L} \subseteq \mathcal{L}'$) then clearly their critical probabilities satisfy $p_c(\mathcal{L}) \geq p_c(\mathcal{L}')$, since any infinite open cluster of $\mathcal{L}$ is contained in some infinite open cluster of $\mathcal{L}'$. When does the *strict* inequality $p_c(\mathcal{L}) > p_c(\mathcal{L}')$ hold? The question may be quantified by asking for non-trivial lower bounds for $p_c(\mathcal{L}) - p_c(\mathcal{L}')$.

Similar questions arise in many ways, not simply within percolation theory. More generally, consider any process indexed by a continuously varying parameter T and experiencing a phase transition at some critical point $T = T_c$. In many cases of interest, sufficient structure is available to enable the conclusion that certain systematic changes to the process can only change T_c in one particular direction. For example, one may be able to conclude that the critical value of the altered process is no greater than T_c. The question then is to understand which systematic changes decrease T_c *strictly*. In the context of the previous paragraph, the systematic changes in question may involve the 'switching on' of edges lying in $\mathcal{L}'$ but not in $\mathcal{L}$.

A related percolation question is that of 'entanglements'. Consider bond percolation on $\mathbb{L}^3$, and examine the box $B(n)$. We think of the open edges as being solid connections made of elastic, say, and we may try to 'pull apart' a pair of opposite faces of $B(n)$. If $p > p_c$, we will generally fail because, with large probability (tending to 1 as $n \to \infty$), there exists an open path joining one face to the opposite face. We may fail even if $p < p_c$, owing to an 'entanglement' of open paths (a chain of necklaces, perhaps, see Figure 3.3). It may be seen that there exists an 'entanglement transition' at some critical point p_c^{ent} satisfying $p_c^{\text{ent}} \leq p_c$. Is it the case that strict inequality holds, that is, that $p_c^{\text{ent}} < p_c$?

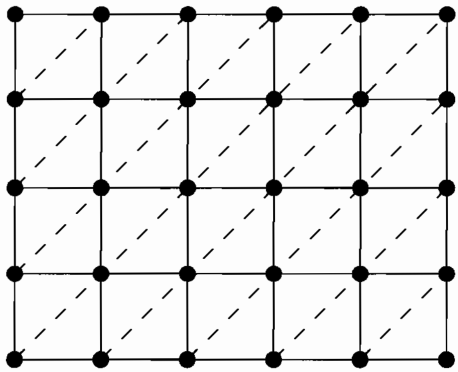


Figure 3.4. The triangular lattice may be obtained from the square lattice by the addition of certain diagonals.

A technology has been developed for approaching such questions of strict inequality. Although, in particular cases, ad hoc arguments can be successful, there appears to be only one general approach known currently. We illustrate this approach next, by sketching the details in a particular case. A more general argument will be presented in Section 3.3, and this will allow an answer to the entanglement question above.

The triangular lattice $\mathbb{T}$ may be obtained by adding diagonals across the squares of the square lattice $\mathbb{L}^2$, in the manner of Figure 3.4. Since any infinite open cluster of $\mathbb{L}^2$ is contained in an infinite open cluster of $\mathbb{T}$, it follows that $p_c(\mathbb{T}) \leq p_c(\mathbb{L}^2)$, but does strict inequality hold? There are various ways of showing that the answer is affirmative. Here we adopt the canonical argument of Aizenman and Grimmett (1991), as an illustration of a general technique. The reason for including this special case in advance of the more general formulation of Theorem (3.16) is that it illustrates clearly the structure of the method with a minimum of complications.

We point out that, *for this particular case*, there is a variety of ways of obtaining the result, by using special properties of the square and triangular lattices. The attraction of the method described below is its generality, relying as it does on essentially no assumptions about lattice structure or number of dimensions.

First we embed the problem in a two-parameter system. Let $p, s \in [0, 1]^2$. We declare each edge of $\mathbb{L}^2$ to be open with probability p, and each *further edge* of $\mathbb{T}$ (that is, the dashed edges in Figure 3.4) to be open with probability s. Writing $P_{p,s}$ for the associated measure, define

$$\theta(p, s) = P_{p,s}(0 \leftrightarrow \infty). \tag{3.3}$$

We propose to establish certain differential inequalities which will imply that $\partial\theta/\partial p$ and $\partial\theta/\partial s$ are comparable, uniformly on any closed subset of the interior $(0, 1)^2$ of the parameter space. This cannot itself be literally achieved, since we have insufficient information about the differentiability of θ. Therefore, we shall

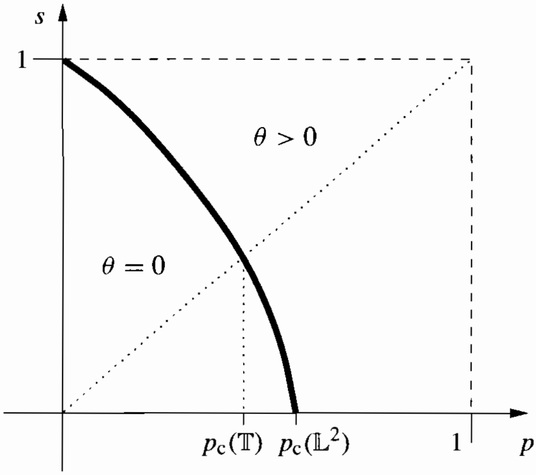


Figure 3.5. The 'critical curve'. The area beneath the curve is the set of pairs (p, s) for which $\theta(p, s) = 0$.

approximate θ by a finite-volume quantity θ_n, and shall work with the partial derivatives of θ_n.

Let $B(n) = [-n, n]^d$, and define

$$\theta_n(p, s) = P_{p,s}\big(0 \leftrightarrow \partial B(n)\big). \tag{3.4}$$

Note that θ_n is a polynomial in p and s, and that

$$\theta_n(p, s) \downarrow \theta(p, s) \qquad \text{as } n \to \infty.$$

(3.5) Lemma. *There exist a positive integer L and a continuous function α mapping $(0, 1)^2$ to $(0, \infty)$ such that*

$$\alpha(p, s)^{-1}\frac{\partial}{\partial p}\theta_n(p, s) \geq \frac{\partial}{\partial s}\theta_n(p, s) \geq \alpha(p, s)\frac{\partial}{\partial p}\theta_n(p, s) \tag{3.6}$$

for $0 < p, s < 1$ *and* $n \geq L$.

Once this is proved, the main result follows immediately, namely the following.

(3.7) Theorem. *It is the case that* $p_c(\mathbb{T}) < p_c(\mathbb{L}^2)$.

Proof of Theorem (3.7). Here is a rough argument, the rigour comes later. It may be shown that there exists a 'critical curve' in (p, s)-space, separating the regime where $\theta(p, s) = 0$ from that when $\theta(p, s) > 0$ (see Figure 3.5). Suppose that this critical curve may be written in the form $h(p, s) = 0$ for some increasing and continuously differentiable function h satisfying $h(p,s) = \theta(p,s)$ whenever $\theta(p,s) > 0$. It is enough to prove that the critical curve contains no vertical segment, and we shall prove this by working with the gradient vector

$$\nabla h = \left(\frac{\partial h}{\partial p}, \frac{\partial h}{\partial s}\right).$$

We take some liberties with (3.6) in the limit as $n \to \infty$, and deduce that

$$\nabla h \cdot (0,1) = \frac{\partial h}{\partial s} \geq \alpha(p,s)\frac{\partial h}{\partial p},$$

whence

$$\frac{1}{|\nabla h|}\frac{\partial h}{\partial s} = \left\{\left(\frac{\partial h}{\partial p}\bigg/\frac{\partial h}{\partial s}\right)^2 + 1\right\}^{-\frac{1}{2}} \geq \frac{\alpha}{\sqrt{\alpha^2+1}},$$

which is bounded away from 0 on any closed subset of $(0,1)^2$. This indicates as required that the critical curve has no vertical segment.

Here is the proper argument. Let η be positive and small, and find γ (> 0) such that $\alpha(p,s) \geq \gamma$ on $[\eta, 1-\eta]^2$. Let $\psi \in [0, \pi/2)$ satisfy $\tan\psi = \gamma^{-1}$.

At the point $(p,s) \in [\eta, 1-\eta]^2$, the rate of change of $\theta_n(p,s)$ in the direction $(\cos\psi, -\sin\psi)$ satisfies

$$\text{(3.8)} \qquad \begin{aligned} \nabla\theta_n \cdot (\cos\psi, -\sin\psi) &= \frac{\partial\theta_n}{\partial p}\cos\psi - \frac{\partial\theta_n}{\partial s}\sin\psi \\ &\leq \frac{\partial\theta_n}{\partial p}(\cos\psi - \gamma\sin\psi) = 0 \end{aligned}$$

by (3.6), since $\tan\psi = \gamma^{-1}$.

Suppose now that $(a,b) \in [2\eta, 1-2\eta]^2$. Let

$$(a',b') = (a,b) + \eta(\cos\psi, -\sin\psi),$$

noting that $(a',b') \in [\eta, 1-\eta]^2$. By integrating (3.8) along the line segment joining (a,b) to (a',b'), we obtain that

$$\text{(3.9)} \qquad \theta(a',b') = \lim_{n\to\infty}\theta_n(a',b') \leq \lim_{n\to\infty}\theta_n(a,b) = \theta(a,b).$$

There is quite a lot of information in such a calculation, but we abstract a small amount only. Let η be small and positive. Take $(a,b) = (p_c(\mathbb{T}) - \zeta, p_c(\mathbb{T}) - \zeta)$ and define (a',b') as above. We choose ζ sufficiently small that $(a,b), (a',b') \in [2\eta, 1-2\eta]^2$, and that $a' > p_c(\mathbb{T})$. The above calculation implies that

$$\text{(3.10)} \qquad \theta(a',0) \leq \theta(a',b') \leq \theta(a,b) = 0,$$

whence $p_c(\mathbb{L}^2) \geq a' > p_c(\mathbb{T})$. □

Proof of Lemma (3.5). With $\mathbb{E}^2$ the edge set of $\mathbb{L}^2$, and $\mathbb{F}$ the additional edges in the triangular lattice $\mathbb{T}$ (that is, the diagonals in Figure 3.4), we have by Russo's formula (in a slightly more general version than Theorem (2.25)) that

$$\begin{aligned} \frac{\partial}{\partial p}\,\theta_n(p,s) &= \sum_{e\in\mathbb{E}^2} P_{p,s}(e \text{ is pivotal for } A_n), \\ \frac{\partial}{\partial s}\,\theta_n(p,s) &= \sum_{f\in\mathbb{F}} P_{p,s}(f \text{ is pivotal for } A_n), \end{aligned} \tag{3.11}$$

where $A_n = \{0 \leftrightarrow \partial B(n)\}$. The idea now is to show that each summand in the first summation is comparable with some given summand in the second. Actually we shall only prove the second inequality in (3.6), since this is the only one used in proving the above theorem, and in addition the proof of the other part is similar.

With each edge e of $\mathbb{E}^2$ we associate a unique edge $f = f(e)$ of $\mathbb{F}$ such that f lies near to e. This may be done in a variety of ways, but in order to be concrete we specify that if $e = \langle z, z+u_1\rangle$ or $e = \langle z, z+u_2\rangle$ then $f = \langle z, z+u_1+u_2\rangle$, where u_1 and u_2 are unit vectors in the directions of the (increasing) x and y axes respectively.

We claim that there exists a function $\beta(p,s)$, continuous on $(0,1)^2$, such that, for all sufficiently large n,

$$P_{p,s}\big(e \text{ is pivotal for } A_n\big) \leq \beta(p,s) P_{p,s}\big(f(e) \text{ is pivotal for } A_n\big) \tag{3.12}$$

for all e lying in $B(n)$. Once this is shown, we sum over e to obtain by (3.11) that

$$\begin{aligned} \frac{\partial}{\partial p}\,\theta_n(p,s) &\leq \beta(p,s) \sum_{e\in\mathbb{E}^2} P_{p,s}\big(f(e) \text{ is pivotal for } A_n\big) \\ &\leq 2\beta(p,s) \sum_{f\in\mathbb{F}} P_{p,s}(f \text{ is pivotal for } A_n) \\ &= 2\beta(p,s) \frac{\partial}{\partial s}\,\theta_n(p,s) \end{aligned}$$

as required. The factor 2 arises because, for each f ($\in \mathbb{F}$), there are exactly two edges e with $f(e) = f$.

The idea of the proof of (3.12) is that, if e is pivotal for A_n in the configuration ω, then, by making 'local changes' to ω, we may create a configuration in which $f(e)$ is pivotal for A_n. The factor β in (3.12) reflects the cost of making such a local change.

Here is a fairly formal proof of (3.12). Suppose that $e = \langle z, z+u_1\rangle$ where $u_1 = (1,0)$; a similar argument will be valid with u_1 replaced by $u_2 = (0,1)$. Let $B_e = z + B(2)$, a box centred at z, and let $\mathbb{E}_e$ be the set of edges of $\mathbb{T}$ having at

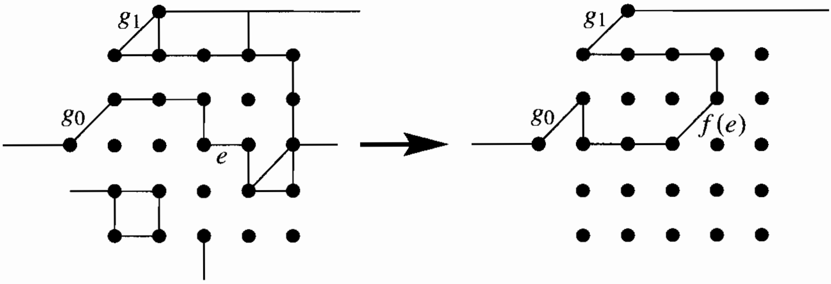


Figure 3.6. An example of a configuration ω on $\mathbb{E}_e$ which gives rise to a configuration $\omega' = \omega'(e, \omega)$.

least one vertex in B_e. Suppose for the moment that $0 \notin B_e$ and $B_e \cap \partial B(n) = \varnothing$. Let ω be a configuration in which e is pivotal for A_n. If e is open, then all paths from 0 to $\partial B(n)$ pass along e. Therefore, there exist two edges $g_i = \langle a_i, b_i \rangle$ of $\mathbb{T}$ (for $i = 1, 2$) such that:

(i) $a_i \in \partial B_e$, $b_i \notin B_e$, and the edge $\langle a_i, b_i \rangle$ is open, for $i = 1, 2$;

(ii) in the configuration obtained from ω by declaring all edges in $\mathbb{E}_e \setminus \{g_0, g_1\}$ to be closed, we have that $0 \leftrightarrow a_0$ and $\partial B(n) \leftrightarrow a_1$.

If there is a choice for the edges g_i then we pick them according to some predetermined ordering of all edges. See Figure 3.6.

Having found the g_i, we may find a configuration ω' ($\in \Omega$) such that:

(iii) ω and ω' agree off $\mathbb{E}_e \setminus \{g_0, g_1\}$;

(iv) $\omega' \in \{f(e)$ is pivotal for $A_n\}$.

The idea for the construction of ω' is to find two vertex-disjoint paths π_0 and π_1 of $\mathbb{T}$ having vertices in B_e, and such that π_0 joins a_0 to z, and π_1 joins a_1 to $z + u_1 + u_2$; then we define ω' by

$$\omega'(h) = \begin{cases} \omega(h) & \text{if } h \notin \mathbb{E}_e \setminus \{g_0, g_1\}, \\ 1 & \text{if } h \text{ lies in } \pi_0 \text{ or } \pi_1, \\ 1 & \text{if } h = f(e), \\ 0 & \text{otherwise.} \end{cases}$$

This construction is illustrated further in Figure 3.7. It may be seen from the figures that ω' satisfies (iv) above. We write $\omega' = \omega'(e, \omega)$ to emphasize the dependence of ω' on the choice of e and ω.

If e is such that either $0 \in B_e$ or $B_e \cap \partial B(n) \neq \varnothing$, then one may find a configuration ω' satisfying (iii) and (iv), although a slightly different geometrical construction is needed for these special cases. We omit the details of this, noting only the conclusion that, for each e and $\omega \in \{e$ is pivotal for $A_n\}$, there exists ω' satisfying (iii) and (iv) above. It follows from (iii) that

$$P_{p,s}(\omega) \leq \frac{1}{\gamma^R} P_{p,s}(\omega')$$

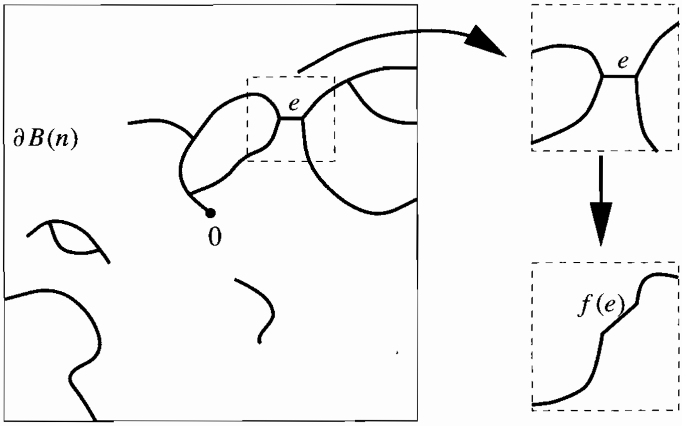


Figure 3.7. Inside the box $B(n)$, the edge e is pivotal for the event $\{0 \leftrightarrow \partial B(n)\}$. By altering the configuration inside the smaller box, we may construct a configuration in which $f(e)$ is pivotal.

where $\gamma = \min\{p, s, 1-p, 1-s\}$ and $R = |\mathbb{E}_e|$.

Write E_h for the event that the edge h is pivotal for A_n. For $\omega \in E_e$, we have by (iv) that $\omega' \in E_{f(e)}$. Therefore,

$$P_{p,s}(E_e) = \sum_{\omega \in E_e} P_{p,s}(\omega) \le \sum_{\omega \in E_e} \frac{1}{\gamma^R} P_{p,s}(\omega') \le \left(\frac{2}{\gamma}\right)^R P_{p,s}(E_{f(e)}),$$

and (3.12) follows with $\beta(p,s) = (2/\gamma)^R$. □

3.3 Enhancements

An 'enhancement' is defined loosely to be a systematic addition of connections according to local rules. Enhancements may involve further coin flips. Can an enhancement create an infinite cluster when previously there was none?

Clearly the answer can be negative. For example, the enhancement may be of the type: join any two neighbours of $\mathbb{L}^d$ with probability $\frac{1}{2}p_c$ whenever they have no incident open edges. Such an enhancement creates extra connections but creates (almost surely) no extra infinite cluster.

Here is a proper definition of the concept of enhancement for bond percolation on $\mathbb{L}^d$ with parameter p. Let R be a positive integer, and let $\mathcal{G}$ be the set of all simple graphs on the vertex set $B = B(R)$. Note that the set of open edges of any configuration ω ($\in \Omega$) generates a member of $\mathcal{G}$, denoted ω_B; $\mathcal{G}$ contains in addition many graphs not obtainable in this way. Let F be a function which associates with every ω_B a graph in $\mathcal{G}$. We call R the 'enhancement range' and F the 'enhancement function'. In the remainder of this chapter, we denote by $e + x$ the translate of an edge e by the vector x; similarly, $G + x$ denotes the translate by x of the graph G on the vertex set $\mathbb{Z}^d$.

We shall consider making an enhancement at each vertex x of $\mathbb{L}^d$, and we shall do this in a stochastic fashion. To this end, we provide ourselves with a vector $\eta = (\eta(x) : x \in \mathbb{Z}^d)$ lying in the space $\Xi = \{0, 1\}^{\mathbb{Z}^d}$. We shall interpret the value $\eta(x) = 1$ as meaning that the enhancement at the vertex x is 'activated'.

These ideas are applied in the following way. For each $x \in \mathbb{Z}^d$, we observe the configuration ω on the box $x + B$, and we write $F(x, \omega)$ for the associated evaluation of F. That is to say, we set $F(x, \omega) = F((\tau_x\omega)_B)$ where τ_x is the shift operator on Ω given by $\tau_x\omega(e) = \omega(e + x)$. The enhanced configuration is defined to be the graph

(3.13) $$G^{\text{enh}}(\omega, \eta) = G(\omega) \cup \left\{ \bigcup_{x:\eta(x)=1} \{x + F(x, \omega)\} \right\}$$

where $G(\omega)$ is the graph of open edges under ω. In writing the union of graphs, we mean the graph with vertex set $\mathbb{Z}^d$ having the union of the appropriate edge sets; wherever this union contains two or more edges between the same pair of vertices, these edges are allowed to coalesce into a single edge.

Thus we associate with each pair $(\omega, \eta) \in \Omega \times \Xi$ an enhanced graph $G^{\text{enh}}(\omega, \eta)$. We endow the sample space $\Omega \times \Xi$ with the product probability measure $P_{p,s}$, and we refer to the parameter s as the *density* of the enhancement.

We call the enhancement function F *essential* if there exists a configuration ω ($\in \Omega$) such that $G(\omega) \cup F(\omega)$ contains a doubly-infinite path but $G(\omega)$ contains no such path. Here are two examples of this definition.

(i) Suppose that F has the effect of adding an edge joining the origin and any given unit vector whenever these two vertices are isolated in $G(\omega)$. In this case, F is not essential.

(ii) If, on the other hand, F adds such an edge whether or not the endvertices are isolated, then F is indeed essential.

We call the enhancement function F *monotonic* if, for all η and all $\omega \le \omega'$, the graph $G^{\text{enh}}(\omega, \eta)$ is a subgraph of $G^{\text{enh}}(\omega', \eta)$. For F to be monotonic it suffices that $\omega_B \cup F(\omega_B)$ be a subgraph of $\omega'_B \cup F(\omega'_B)$ whenever $\omega \le \omega'$.

The *enhanced percolation probability* is defined as

(3.14) $$\theta^{\text{enh}}(p, s) = P_{p,s}\big(0 \text{ belongs to an infinite cluster of } G^{\text{enh}}\big).$$

A useful definition of the *enhancement critical point* is given by

(3.15) $$p_c^{\text{enh}}(F, s) = \inf\{p : \theta^{\text{enh}}(p, s) > 0\}.$$

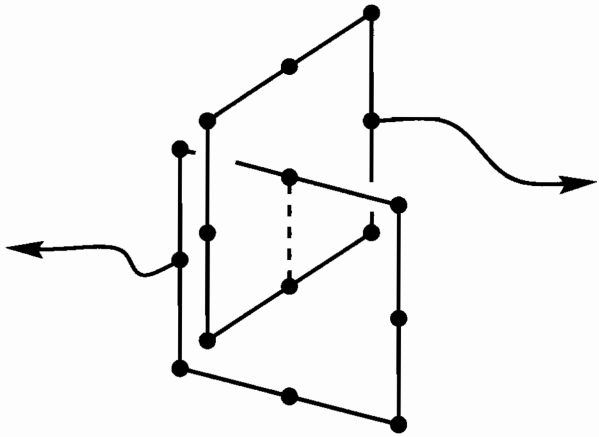


Figure 3.8. A sketch of the enhancement which adds an edge between any two interlocking 2×2 squares in $\mathbb{L}^3$. This picture contains a doubly-infinite path if and only if the enhancement is activated.

We note from (3.13) that θ^{enh} is non-decreasing in s. If F is monotonic then, by Theorem (2.1), θ^{enh} is non-decreasing in p also, whence

$$\theta^{\text{enh}}(p, s) \begin{cases} = 0 & \text{if } p < p_{\text{c}}^{\text{enh}}(F, s), \\ > 0 & \text{if } p > p_{\text{c}}^{\text{enh}}(F, s). \end{cases}$$

If F is not monotonic, there will generally be ambiguity over the correct definition of the critical point. We will abide by (3.15) here.

(3.16) Theorem. *Let $s > 0$. If the enhancement function F is essential, then $p_{\text{c}}^{\text{enh}}(F, s) < p_{\text{c}}$.*

That is to say, essential enhancements shift the critical point *strictly*. Instead of enhancements, one may study 'diminishments', which involve the systematic removal of open edges according to some local rule. A similar theorem may then be formulated, asserting that the critical point is strictly increased so long as the diminishment in question satisfies a condition parallel to that of essentialness given above.

Here are some examples of Theorem (3.16) and related arguments.

A. *Entanglements.* Consider bond percolation on the three-dimensional cubic lattice $\mathbb{L}^3$. Whenever we see two interlinking 2×2 open squares, we join them by an edge (see Figure 3.8). It is easy to see that this enhancement is essential, and therefore it shifts the critical point downwards. Any reasonable definition of entanglement would require that two such interlocking squares be entangled, and it would follow that $p_{\text{c}}^{\text{ent}} < p_{\text{c}}$. We do not formulate precisely the notion of an entanglement since there are certain difficulties over this; see Section 12.5.

B. *Lattices and sublattices.* Let $\mathcal{L}$ be a sublattice of the lattice $\mathcal{L}'$. Assuming a reasonable definition of the term 'lattice', there will exist a periodic class $\mathcal{E}$ of edges of $\mathcal{L}'$ which do not lie in $\mathcal{L}$. Suppose it is the case that each $e \in \mathcal{E}$ is such that: there exists a bond configuration on $\mathcal{L}$ containing no doubly-infinite open path, but such a path exists if we add e to the configuration. Although Theorem (3.16) cannot be applied directly in this situation, its proof may be adapted in a straightforward manner to deduce (rather as in Section 3.2) that $p_c(\mathcal{L}) > p_c(\mathcal{L}')$.

C. *Slabs.* Let $d \geq 3$ and $k \geq 0$, and define the slab S_k of thickness k by $S_k = \mathbb{Z}^2 \times \{0, 1, 2, \dots, k\}^{d-2}$. Since $S_k \subseteq S_{k+1}$, we have that $p_c(S_k) \geq p_c(S_{k+1})$. The method of Theorem (3.16) may be used as follows to obtain the strict inequality $p_c(S_k) > p_c(S_{k+1})$. Let e be the unit vector $(0, 0, \dots, 0, 0, 1)$. Take $\mathcal{L}'$ to be the graph derived from $\mathbb{L}^d$ by deleting all edges of the form $\langle x, x+e \rangle$ such that $|x_d + 1|$ is divisible by $k+2$. We construct the subgraph $\mathcal{L}$ of $\mathcal{L}'$ by deleting all edges $\langle x, x+e \rangle$ such that $|x_d + 1|$ is divisible by $k+1$. Then $\mathcal{L}'$ may be obtained by systematic enhancements of $\mathcal{L}$, and the claim may now be obtained in the usual way.

D. *Augmented percolation.* Here is a question which has arisen in so called 'invasion percolation'. Consider bond percolation on a lattice $\mathcal{L}$. Each edge is in exactly one of three categories: (i) open, (ii) closed and belonging to a finite closed cluster, (iii) closed and belonging to an infinite closed cluster. Consider the graph obtained from $\mathcal{L}$ by deleting all edges lying in category (iii) while retaining those in categories (i) and (ii). Does there exist an interval of values of p ($< p_c$) for which this graph contains (almost surely) an infinite component? That this indeed holds for $\mathbb{L}^d$ with $d \geq 2$ follows by considering the enhancement in which an edge is added between the origin and a neighbour x if and only if all other edges incident to 0 and x are open.

E. *Site percolation.* The proof of Theorem (3.16) may be adapted to bond and site percolation on general lattices. The condition of 'essentialness' was formulated above for bond percolation, and is replaced as follows for site percolation. We say that the realization $\xi \in \{0, 1\}^{\mathbb{Z}^d}$ of site percolation contains a doubly-infinite *self-repelling* path if there exists a doubly-infinite open path none of whose vertices is adjacent to any other vertex of the path except for its two neighbours in the path. An enhancement of site percolation is called *essential* if there exists a configuration ξ containing no doubly-infinite self-repelling path, but such that the enhanced configuration obtained by activating the enhancement at the origin does indeed contain such a path.

Proof of Theorem (3.16). We follow Aizenman and Grimmett (1991). In this proof we shall construct various functions on the parameter space $(0, 1)^2$, denoted as $\delta_i = \delta_i(p, s)$ for $i \geq 1$. Such functions shall by convention be continuous and strictly positive on their domain $(0, 1)^2$.

The first step is to generalize equations (3.11). A pair $(\omega, \eta) \in \Omega \times \Xi$ gives rise to an enhanced graph $G^{\text{enh}}(\omega, \eta)$, and we call the edges of this graph *enhanced*. For $(\omega, \eta) \in \Omega \times \Xi$ and $e \in \mathbb{E}^d$, $x \in \mathbb{Z}^d$, we define configurations ω^e, ω_e, η^x, η_x by

$$\omega^e(f) = \begin{cases} \omega(f) & \text{if } f \neq e, \\ 1 & \text{if } f = e, \end{cases} \qquad \omega_e(f) = \begin{cases} \omega(f) & \text{if } f \neq e, \\ 0 & \text{if } f = e, \end{cases}$$

$$\eta^x(y) = \begin{cases} \eta(y) & \text{if } y \neq x, \\ 1 & \text{if } y = x, \end{cases} \qquad \eta_x(y) = \begin{cases} \eta(y) & \text{if } y \neq x, \\ 0 & \text{if } y = x. \end{cases}$$

Let n be a positive integer, and let $A = A_n$ be the event that there exists a path of enhanced edges joining the origin to some vertex of the set $\partial B(n)$. For $(\omega, \eta) \in \Omega \times \Xi$ and $e \in \mathbb{E}^d$, $x \in \mathbb{Z}^d$, we say that

e is (+)pivotal for A if $I_A(\omega^e, \eta) = 1$ and $I_A(\omega_e, \eta) = 0$,
e is (−)pivotal for A if $I_A(\omega^e, \eta) = 0$ and $I_A(\omega_e, \eta) = 1$,
x is (+)pivotal for A if $I_A(\omega, \eta^x) = 1$ and $I_A(\omega, \eta_x) = 0$,

where

$$I_A(\omega, \eta) = \begin{cases} 1 & \text{if } A \text{ occurs}, \\ 0 & \text{otherwise}. \end{cases}$$

Note that, if the enhancement is not monotonic, edges may generally be either (−)pivotal or (+)pivotal for an increasing event A. Vertices, on the other hand, can only be (+)pivotal for an increasing event.

Since the occurrence of A depends on only finitely many of the $\omega(e)$ and $\eta(x)$, we have by a minor extension of Theorem (2.32) that $\theta_n(p, s) = P_{p,s}(A)$ satisfies

(3.17)
$$\frac{\partial \theta_n}{\partial p} = \sum_{e \in \mathbb{E}^d} \Big\{ P_{p,s}\big(e \text{ is (+)pivotal for } A\big) - P_{p,s}\big(e \text{ is (−)pivotal for } A\big) \Big\},$$

$$\frac{\partial \theta_n}{\partial s} = \sum_{x \in \mathbb{Z}^d} P_{p,s}\big(x \text{ is (+)pivotal for } A\big).$$

We continue with a geometrical observation. Recall that R is the range of the enhancement. Let m be a positive integer satisfying $m > R + 2$, and let v, w be distinct vertices in $\partial B(m)$. The enhancement has been assumed essential, which is to say that there exists a bond configuration ω having the following property: ω contains no doubly-infinite open path, but such a path $\pi = \pi(\omega)$ is created when the enhancement at the origin is activated. Such a path π must contain two disjoint singly-infinite open paths of ω, denoted $\pi_1 = x_0, f_0, x_1, f_1, \ldots$ and $\pi_2 = y_0, g_0, y_1, g_1, \ldots$, such that $x_0, y_0 \in B(R)$. Let

$$r = \min\big\{i : x_i \in \partial B(m-1)\big\}, \quad s = \min\big\{i : y_i \in \partial B(m-1)\big\},$$

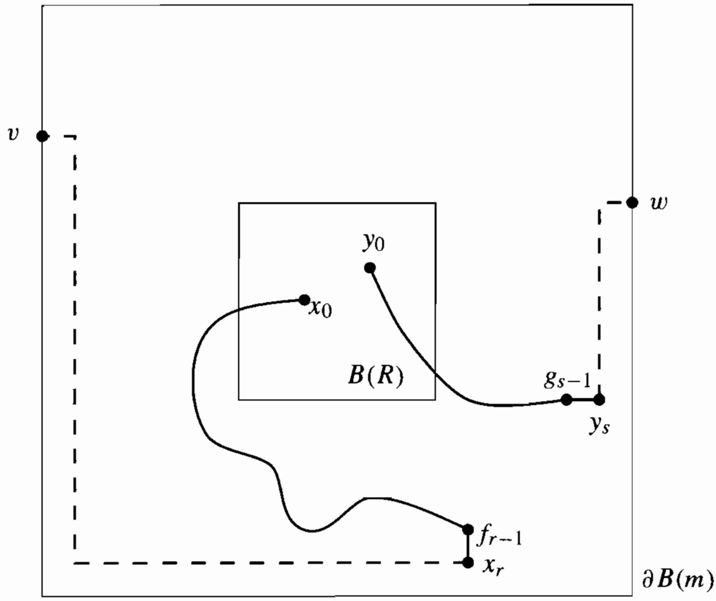


Figure 3.9. An illustration of the way in which $\widehat{\omega}$ is constructed when the enhancement is essential.

noting that $r, s \geq 2$. We may find vertex-disjoint paths ν_1, ν_2 of $\mathbb{L}^d$, using vertices in $B(m) \setminus B(m-2)$ only, such that ν_1 joins x_r to v and ν_2 joins y_s to w. We now define the configuration $\widehat{\omega}$ ($\in \Omega$) by:

$$\widehat{\omega}(e) = \begin{cases} \omega(e) & \text{if } e \text{ has both its endvertices in } B(m-2), \\ 1 & \text{if } e \text{ lies in either } \nu_1 \text{ or } \nu_2, \\ 1 & \text{if } e = f_{r-1}, g_{s-1}, \\ 0 & \text{otherwise.} \end{cases}$$

About $\widehat{\omega}$ we note the following:

(i) all open edges of $\widehat{\omega}$ have both endvertices in $B(m)$;

(ii) $\widehat{\omega}$ has no open path joining v and w;

(iii) if the enhancement at 0 is activated; then an enhanced path is created joining v and w and using vertices of $B(m)$ only.

We write $\widehat{\omega} = \widehat{\omega}_m(v, w)$ to emphasize the dependence of $\widehat{\omega}$ on n, v, w. The construction of $\widehat{\omega}$ is illustrated in Figure 3.9.

Suppose that $e = \langle z, z+u \rangle$ where u is a (positive) unit vector of $\mathbb{L}^d$. Let $B_e = z + B(m+R)$, a box centred at z, and let $v_1, v_2, \ldots$ be a fixed ordering of the vertices of B_e. For $\eta \in \Xi$, we define η_i by

$$\eta_i(x) = \begin{cases} 0 & \text{if } x \in \{v_1, v_2, \ldots, v_i\}, \\ \eta(x) & \text{otherwise.} \end{cases} \tag{3.18}$$

Let n be a positive integer, write $A = A_n$, and let $(\omega, \eta) \in \Omega \times \Xi$. Suppose for the moment that

$$m + 1 \leq \|z\| \leq n - m - 1, \tag{3.19}$$

and let

$$K_e = \min\big\{i : \text{ some vertex of } B_e \text{ is (+)pivotal for } A \text{ in the configuration } (\omega, \eta_i)\big\},$$

with the convention that the minimum of the empty set is ∞.

The configurations (ω, η_i) are obtained from (ω, η) by altering a bounded number of variables $\eta(x)$. Also, if $K_e < \infty$, then in at least one of the configurations (ω, η_i), for $0 \leq i \leq |B_e|$, there exist one or more (+)pivotal vertices. Therefore, there exists a function δ_1 such that

(3.20)

$$\begin{aligned} P_{p,s}(K_e < \infty) &\leq \sum_{i=0}^{|B_e|} \sum_{x \in B_e} P_{p,s}\big(\{(\omega, \eta) : x \text{ is (+)pivotal for } A \text{ in } (\omega, \eta_i)\}\big) \\ &\leq \delta_1(p, s)\big(1 + |B_e|\big)^2 P_{p,s}(\Pi_e \geq 1), \end{aligned}$$

where Π_e is the number of (+)pivotal vertices for A lying in B_e; this may be compared with the final step in the proof of Lemma (3.5).

We consider next the case $K_e = \infty$. Let $(\omega, \eta) \in \Omega \times \Xi$ be such that e is (+)pivotal for A, $\omega(e) = 1$, and $K_e = \infty$. Let η' be given by

$$\eta'(x) = \begin{cases} 0 & \text{if } x \in B_e, \\ \eta(x) & \text{otherwise.} \end{cases}$$

Since $K_e = \infty$, we have that e is (+)pivotal for A in (ω, η'). Using (3.19) and the fact that $\omega(e) = 1$, we observe that there exists an enhanced path $x_0 = 0, f_0, x_1, f_1, \dots, x_t$ with $x_t \in \partial B(n)$ which utilizes the edge e, and we set

$$r = \min\big\{i : x_i \in z + B(m)\big\}, \quad s = \max\big\{i : x_i \in z + B(m)\big\},$$

the first and last vertices thereof lying in $z + B(m)$. Note that $1 \leq r < s < t$. Let $\mathbb{E}_e$ be the set of lattice edges having at least one endvertex in $z + B(m)$. We propose to alter the values $\omega(f)$, $f \in \mathbb{E}_e$, in order to obtain a new configuration in which z is (+)pivotal for A. We do this by 'pasting' the configuration $\widehat{\omega} = \widehat{\omega}_m(x_r - z, x_s - z)$ into the box $z + B(m)$. More specifically, we define ω' $(\in \Omega)$ by

$$\omega'(h) = \begin{cases} \widehat{\omega}(h - z) & \text{if } h \text{ has both endvertices in } B_e, \\ 1 & \text{if } h = f_{r-1}, f_s, \\ 0 & \text{for other edges } h \text{ of } \mathbb{E}_e, \\ \omega(h) & \text{if } h \notin \mathbb{E}_e. \end{cases} \tag{3.21}$$

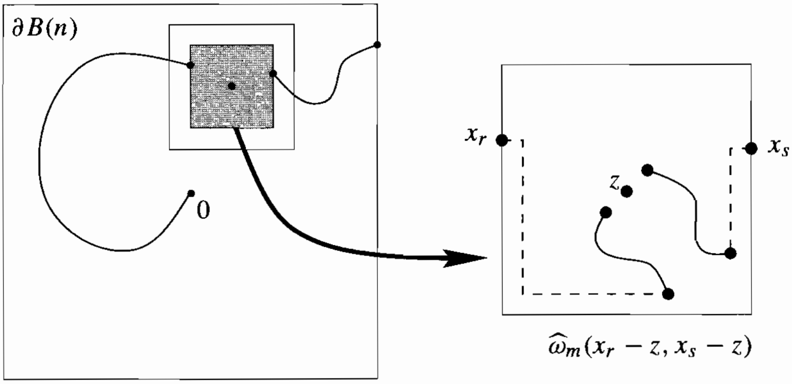


Figure 3.10. If $K_e = \infty$, one may alter the states of a bounded number of edges and vertices in order to obtain a configuration ω' in which z is $(+)$pivotal for A.

The configuration ω' is illustrated in Figure 3.10.

It may be seen from the definition of $\widehat{\omega}$ that, in (ω', η'), the vertex z is $(+)$pivotal for A. Since (ω', η') has been obtained from (ω, η) by changing only a bounded number of variables $\omega(h)$, $\eta(x)$, there exists δ_2 such that

$$
\begin{aligned}
(3.22) \qquad P_{p,s}\big(e \text{ is } (+)\text{pivotal}, e \text{ is open}, K_e = \infty\big) &\le \delta_2(p,s) P_{p,s}\big(z \text{ is } (+)\text{pivotal for } A\big) \\
&\le \delta_2(p,s) P_{p,s}(\Pi_e \ge 1).
\end{aligned}
$$

Adding (3.20) and (3.22), and remembering that the events $\{e \text{ is } (+)\text{pivotal for } A\}$ and $\{e \text{ is open}\}$ are independent, we conclude that

$$
(3.23) \qquad P_{p,s}\big(e \text{ is } (+)\text{pivotal for } A\big) \le \delta_3(p,s) P_{p,s}(\Pi_e \ge 1)
$$

for some δ_3.

We now relax assumption (3.19). Suppose first that $z \in B(m)$ and that e is $(+)$pivotal for A. Instead of working in the box $B_e = z + B(m)$, we work instead within the larger box $B(2m+1)$. If the quantity corresponding to K_e is finite, then the above argument may be applied directly. If it is infinite, we alter the configuration within $B(2m+R+1)$ in such a way as to arrange for the vertex $(m+1, 0, 0, \dots, 0)$ to become $(+)$pivotal for A. This leads as before to an inequality of the form of (3.23) with δ_3 replaced by some δ_4 and with Π_e replaced by the number of $(+)$pivotal vertices inside the box $B(2m+R+1)$.

A similar construction is valid if $\|z\| \ge n - m$, although we note the added complication that there may exist $(+)$pivotal vertices which lie outside $B(n)$, but necessarily within distance R of $B(n)$.

In conclusion, there exists δ_4 such that, for all $e = \langle z, z+u \rangle \in \mathbb{E}^d$,

$$
(3.24) \qquad P_{p,s}\big(e \text{ is } (+)\text{pivotal for } A\big) \le \delta_4(p,s) P_{p,s}(\Pi'_e \ge 1)
$$

where Π'_e is the number of pivotal vertices within $z + B(2m + R + 1)$.

Summing (3.24) over all $e \in \mathbb{E}^d$, we deduce via (3.17) that

$$\frac{\partial \theta_n}{\partial p} \le \delta_4(p,s) d |B(2m+R+1)| \frac{\partial \theta_n}{\partial s} = \nu(p,s)\frac{\partial \theta_n}{\partial s}, \tag{3.25}$$

just as in the second inequality of Lemma (3.5). We now argue as in the proof of (3.8)–(3.9). Let η be positive and small, and choose γ such that $\nu(p,s) \le \gamma^{-1}$ on $[\eta, 1-\eta]^2$, and let $\tan\psi = \gamma^{-1}$. If $(a,b) \in [2\eta, 1-2\eta]^2$ and

$$(a', b') = (a,b) + \eta(\cos\psi, -\sin\psi), \tag{3.26}$$

then, as in (3.9),

$$\theta(a', b') \le \theta(a,b). \tag{3.27}$$

Let $0 < b < 1$ and let η (> 0) be sufficiently small that

$$2\eta < b,\ p_c(\mathbb{L}^d) < 1 - 2\eta.$$

We may find a such that

$$2\eta < a < p_c(\mathbb{L}^d) < a' < 1 - 2\eta$$

where a' is given in (3.26). By (3.27),

$$\theta(a,b) \ge \theta(a',b') \ge \theta(a',0) > 0,$$

whence $p_c^{\text{enh}}(F, b) \le a$ as required. □

3.4 Bond and Site Critical Probabilities

For any connected graph G, it is the case that $p_c^{\text{bond}}(G) \le p_c^{\text{site}}(G)$, but when does strict inequality hold here? The answer depends on the choice of graph. For example, if G is a tree, it is easy to see that *equality* holds rather than *inequality*. On the other hand, it is reasonable to expect strict inequality to be valid for a range of graphs including all finite-dimensional lattices in two or more dimensions. We prove this in the special case of $\mathbb{L}^d$ with $d \ge 2$.

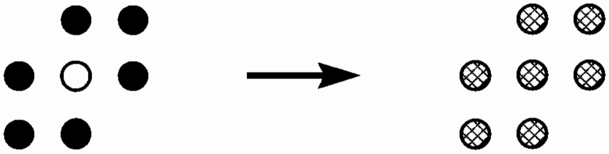


Figure 3.11. An illustration of the enhancement described, when $d = 2$. Each copy of the configuration on the left is replaced, with probability s, by the configuration on the right. Filled circles indicate open vertices, and hatched circles denote enhanced vertices.

(3.28) Theorem. *Consider* $\mathbb{L}^d$ *with* $d \geq 2$. *We have that* $p_c^{\text{bond}} < p_c^{\text{site}}$.

Proof. We follow Grimmett and Stacey (1998). The basic approach is to use the enhancement technology expounded in the last section, but with some interesting differences. We shall construct a site percolation process on $\mathbb{L}^d$, and shall define an enhancement thereof which is dominated by bond percolation.

The sample space appropriate for site percolation is $\Xi = \{0, 1\}^{\mathbb{Z}^d}$. We interpret the vector $\xi \in \Xi$ as a realization of site percolation on $\mathbb{L}^d$. At each vertex x, we shall consider making an enhancement with probability s, and to this end we provide ourselves with an 'enhancement realization' $\eta \in \Xi$. As before, we interpret the value $\eta(x) = 1$ as meaning that the enhancement at the vertex x is activated. The pair (ξ, η) takes values in the sample space $\Xi \times \Xi$, and we endow this space with the product probability measure $P_{p,s} = P_p \times P_s$.

Let $u_1, u_2, \dots, u_d$ denote the unit vectors of $\mathbb{R}^d$, that is, $u_1 = (1, 0, 0, \dots, 0)$, $u_2 = (0, 1, 0, 0, \dots, 0)$, and so on. Given a vertex $x \in \mathbb{Z}^d$ we define disjoint sets of vertices close to x as follows:

$$A_x = \{x + u_i,\ x + u_i + u_j : 1 \leq i < j \leq d\},$$
$$B_x = \{x - u_i,\ x - u_i - u_j : 1 \leq i < j \leq d\}.$$

We say that a vertex x is a *qualifying vertex* (for ξ) if $\xi(x) = 0$ and in addition $\xi(y) = 1$ for all $y \in A_x \cup B_x$. For $(\xi, \eta) \in \Xi \times \Xi$, the enhanced configuration $\zeta = \zeta(\xi, \eta)$ ($\in \Xi$) is defined by: $\zeta(x) = 1$ if and only if either

(i) $\xi(x) = 1$, or

(ii) x is a qualifying vertex for ξ, and $\eta(x) = 1$.

We call a vertex x *open* if $\xi(x) = 1$, and *enhanced* if $\zeta(x) = 1$. See Figure 3.11 for a sketch of the above enhancement in action.

We shall refer to ζ (or the law it induces on Ξ) as *enhanced site percolation with parameters p and s*, and we write

$$\theta^{\text{enh}}(p, s) = P_{p,s}\big(0 \text{ lies in an infinite enhanced path}\big)$$

for the percolation probability of the enhanced configuration; cf. (3.14).

(3.29) Lemma. *We have that* $\theta^{\text{enh}}(p, p^2) \le \theta^{\text{bond}}(p)$.

Theorem (3.28) follows easily from this lemma, as follows. We note that Theorem (3.16) is not directly applicable in this setting, since it was concerned with enhancements of bond percolation rather than of site percolation. However, it is straightforward to adapt the theorem to the current setting, and it may be seen that the enhancement described above is essential in the sense of site percolation; see Paragraph E following the statement of Theorem (3.16). Let s satisfy $\sqrt{s} = \frac{1}{2} p_{\text{c}}^{\text{site}}$. It follows from the appropriate reworking of Theorem (3.16) that there exists $\pi(s)$ satisfying $\pi(s) < p_{\text{c}}^{\text{site}}$ such that $\theta^{\text{enh}}(p, s) > 0$ for all $p > \pi(s)$. Let p satisfy

$$\max\{\pi(s), \sqrt{s}\} < p < p_{\text{c}}^{\text{site}}.$$

Since $p^2 > s$, we have that $\theta^{\text{enh}}(p, p^2) \ge \theta^{\text{enh}}(p, s) > 0$. Therefore, by Lemma (3.29), $\theta^{\text{bond}}(p) > 0$, whence $p_{\text{c}}^{\text{site}} > p \ge p_{\text{c}}^{\text{bond}}$ as required.

Proof of Lemma (3.29). We shall employ a coupling of bond and site percolation which is essentially that used for the second inequality of (1.34). Let $X = (X_e : e \in \mathbb{E}^d)$ be a realization of bond percolation on $\mathbb{L}^d$. Let $Z = (Z_x : x \in \mathbb{Z}^d)$ be a collection of independent Bernoulli random variables, independent of the X_e, having mean p also. In the first stage of this proof, we construct from these two families a new collection $Y = (Y_x : x \in \mathbb{Z}^d)$ of random variables, which constitutes a site percolation process with density p. This last process will have the property that, for $x, y \in \mathbb{Z}^d$, if y cannot be reached from x in the bond process X, then neither can y be reached from x in the site process Y; this will allow the conclusion $\theta^{\text{site}}(p) \le \theta^{\text{bond}}(p)$.

Let $e_1, e_2, \ldots$ be an enumeration of the edges of $\mathbb{Z}^d$ and let $x_1, x_2, \ldots$ be an enumeration of its vertices; we take $x_1 = 0$, the origin. We wish to define the Y_x in terms of the X_e and the Z_y, and we shall do so by a recursion, described next. Suppose at some stage that we have defined the set $(Y_x : x \in S)$, where S is a proper subset of $\mathbb{Z}^d$. (At the start we take $S = \varnothing$.) For $x \in S$, we say that x is 'currently open' if $Y_x = 1$ and 'currently closed' if $Y_x = 0$. Let T be the set of vertices not belonging to S which are adjacent to some currently open vertex. If $T = \varnothing$, then let y be the first vertex (in the above enumeration) not lying in S, and set $Y_y = Z_y$. If $T \ne \varnothing$, we let y be the first vertex in T, and we let y' be the first currently open vertex adjacent to it; we then set $Y_y = X_{\langle y, y' \rangle}$, where as usual $\langle u, v \rangle$ denotes the edge joining two neighbours u, v. Repeating this procedure will eventually exhaust all vertices $x \in \mathbb{Z}^d$, and assign values to all the random variables Y_x.

This algorithm begins at the origin 0, and builds up a (possibly infinite) open cluster together with a neighbour set of closed vertices. When the cluster at 0 is complete, another vertex is selected as a new starting point, and the process is iterated. Note that this recursion is transfinite, since infinitely many steps are needed in order to build up any infinite cluster.

We now make two observations about the variables Y_x. First, for each vertex x, the probability that $Y_x = 1$, conditional on any information about the values of those Y_y determined prior to the definition of Y_x, is equal to p. Based upon this observation one may prove without great difficulty that the random variables $(Y_x : x \in \mathbb{Z}^d)$ are independent with mean p, which is to say that they form a site percolation process on $\mathbb{Z}^d$.

Secondly, if there exists a path of open vertices of Y between two points, then there exists a (possibly longer) path of open bonds of X. Therefore, we have succeeded in coupling a bond and a site process with the required domination property.

We shall now adapt this construction in order to obtain a suitable coupling of bond percolation with the enhanced site percolation process obtained from the Y_x. Here is the main idea. Suppose that x is a qualifying vertex for the realization Y. Then $Y_x = 0$, and $Y_y = 1$ for all $y \in A_x \cup B_x$. Note that all the vertices of A_x (respectively B_x) must lie in the same site percolation cluster $C_1 = C_1(x)$ (respectively $C_2 = C_2(x)$). If $C_1 = C_2$, then the enhancing of x makes no difference to the connectivity properties of the graph except at x. If $C_1 \neq C_2$, then enhancing x effectively joins C_1 and C_2 together. Since $Y_x = 0$, it is the case that at most one edge e incident with x was examined (in the sense that the value of X_e was considered) in the determination of the Y_u. Therefore, there exists at least one unexamined edge joining x to A_x; let the first such edge in our enumeration be $e = e(x)$. Likewise, there exists a first unexamined edge, $f = f(x)$ say, joining x and B_x. We adopt the following rule: we declare x to be enhanced if and only if $X_e = X_f = 1$. This has the effect of adding x into the enhanced configuration with probability p^2. Acting thus for all qualifying vertices x yields an enhanced site percolation; the independence of the enhancement at different qualifying vertices follows from the fact that the sequence of all $e(x)$ and $f(x)$ contains no repetitions. Furthermore, the above enhancement cannot join any two vertices which are not already joined by an open path in the bond model: enhancing x has the effect of connecting x to the clusters $C_1(x)$ and $C_2(x)$ and to no others, and this enhancement of x occurs only in situations where x is already joined to both of these clusters in the bond process X.

It is fairly straightforward to present a formal description of the informal account above. In order to obtain the appropriate enhancement, we require a family $(H_x : x \in \mathbb{Z}^d)$ of independent Bernoulli random variables, having parameter p^2 and independent of the vector Y. We only require the H_x for qualifying vertices x, and we may simply set $H_x = X_{e(x)}X_{f(x)}$, where $e(x)$ and $f(x)$ are given as above.

We have now given a coupling of bond percolation and an enhanced site percolation with the property that any two vertices which are in the same cluster of the enhanced site process are also in the same cluster of the bond process. It follows that, if the origin lies in an infinite enhanced path, then the cluster containing the origin in the bond process is infinite also. The required inequality follows. □

3.5 Notes

Section 3.1. We omit a detailed history of the results of this section, of which a discussion may be found in Hughes (1996). Kesten (1980a, 1982) proved that $p_c = \frac{1}{2}$ for bond percolation on $\mathbb{L}^2$, and Wierman (1981) adapted his proof in order to calculate p_c for the hexagonal and triangular lattices. These rigorous arguments confirmed the proposals of Sykes and Essam (1963, 1964), who discussed the notion of a matching pair of graphs. The exact calculation of p_c(bow-tie) appeared in Wierman (1984a).

Certain rigorous numerical inequalities have been proved for two-dimensional percolation by Wierman (1990, 1995). The rigorous derivation of the series expansion (3.2) was presented by Hara and Slade (1995), in response to physical arguments which appeared earlier in the physics literature.

Sections 3.2 and 3.3. The first systematic approach to strict inequalities for ordered pairs of lattices is due to Menshikov (1987a, d, e), although there existed already some special results in the literature. The discussion and technology of Sections 3.2 and 3.3 draws heavily on Aizenman and Grimmett (1991); see also Grimmett (1997).

Theorem (3.16) may be adapted to enhancements of site percolation (see the discussion following the statement of the theorem). The assumption that enhancements take place at *all* vertices x may be relaxed; see Aizenman and Grimmett (1991).

The problem of entanglements appeared first in Kantor and Hassold (1988), who reported certain numerical conclusions. The existence of an entanglement transition different from that of percolation was proved by Aizenman and Grimmett (1991); the strict positivity of the entanglement critical point was proved by Holroyd (1998b). The entanglement transition has been studied more systematically by Grimmett and Holroyd (1998); in particular, they have discussed certain topological difficulties in deciding on the 'correct' definition of an infinite entanglement and of the entanglement critical point.

Related issues arise in the study of so called 'rigidity percolation', in which one studies the existence of infinite rigid components of the open subgraph of a lattice; see Jacobs and Thorpe (1995, 1996) and Holroyd (1998a). Further accounts of entanglement and rigidity may be found in Sections 12.5 and 12.6.

The 'augmented percolation' question posed after Theorem (3.16) was discussed by Chayes, Chayes, and Newman (1984) in the context of invasion percolation on the triangular lattice and on the covering lattice of the square lattice. See also Pokorny, Newman, and Meiron (1990).The question in its present form was answered by Aizenman and Grimmett (1991).

Section 3.4. Theorem (3.28) is taken from Grimmett and Stacey (1998), where a general theorem of this sort is presented. Earlier work on strict inequalities between bond and site critical probabilities in two dimensions may be found in

Higuchi (1982), Kesten (1982), and Tóth (1985). Corresponding results for Ising, Potts, and random-cluster models have been studied by Aizenman and Grimmett (1991), Bezuidenhout, Grimmett, and Kesten (1993), and by Grimmett (1993, 1994a, 1995a, 1999c).

Chapter 4

The Number of Open Clusters per Vertex

4.1 Definition

In addition to the percolation probability $\theta(p)$ and the mean cluster size $\chi(p)$, there is a third principal character in percolation theory. The *number of open clusters per vertex* is defined as in (1.30) by

$$\kappa(p) = E_p(|C|^{-1}) = \sum_{n=1}^{\infty} \frac{1}{n} P_p(|C| = n), \tag{4.1}$$

where $|C|$ is the number of vertices in the open cluster at the origin. The function κ crops up frequently but usually peripherally in calculations: for example, it was a central figure in the first proof that the infinite open cluster is (almost surely) unique when $p > p_c$ (see Section 8.2), as well as in the famous Sykes–Essam 'proof' that $p_c = \frac{1}{2}$ for bond percolation on $\mathbb{L}^2$ (see Theorem (11.4)). We begin by exploring the origins of κ and its rather unwieldy title.

Let $B(n)$ be the box with side-length $2n$ and centre at the origin, as usual. We think of $B(n)$ as being a graph, by adding the edges inherited from $\mathbb{L}^d$. The principal object of current attention is the number K_n of open clusters of $B(n)$. More precisely, K_n is the number of connected components obtained in $B(n)$ by deleting all closed edges. It turns out that K_n is approximately a linear function of the volume of $B(n)$, in that $K_n|B(n)|^{-1}$ converges to a non-trivial limit as $n \to \infty$.

(4.2) Theorem. *Suppose $0 \le p \le 1$. The number K_n of open clusters of $B(n)$ satisfies*

$$\frac{1}{|B(n)|} K_n \to \kappa(p) \qquad \text{as } n \to \infty,\ P_p\text{-a.s. and in } L^1(P_p). \tag{4.3}$$

This explains the title of the function κ. The idea of the proof is simple. For each $x \in \mathbb{Z}^d$, we define

$$\Gamma(x) = \begin{cases} |C(x)|^{-1} & \text{if } |C(x)| < \infty, \\ 0 & \text{if } |C(x)| = \infty. \end{cases} \tag{4.4}$$

Then $\sum_{x \in B(n)} \Gamma(x)$ is approximately equal to K_n, since any open cluster of $\mathbb{L}^d$ which is contained entirely in $B(n)$ contributes exactly 1 to this sum. There are certain 'boundary effects' which arise from points x in open clusters not contained entirely in $B(n)$, but these turn out to be negligible when we take the limit:

$$\lim_{n\to\infty} \left\{ \frac{1}{|B(n)|} \sum_{x \in B(n)} \Gamma(x) \right\} = \lim_{n\to\infty} \left\{ \frac{1}{|B(n)|} K_n \right\}. \tag{4.5}$$

Finally, we use the ergodic theorem to see that the first limit equals $E_p(\Gamma(0)) = E_p(|C|^{-1})$. Here is a formal proof.

Proof of Theorem (4.2). For $x \in B(n)$ we define $C_n(x)$ to be the open cluster of $B(n)$ containing x; that is to say, we delete all closed edges from $B(n)$, and we write $C_n(x)$ for the connected component of the ensuing graph which contains the vertex x. As usual, we write $|C_n(x)|$ for the number of vertices of $C_n(x)$.

For each $x \in \mathbb{Z}^d$, we define $\Gamma(x) = |C(x)|^{-1}$. For $x \in B(n)$, we define $\Gamma_n(x) = |C_n(x)|^{-1}$, noting that

$$\Gamma(x) \le \Gamma_n(x) \qquad \text{for all } x \in B(n). \tag{4.6}$$

The basic formula of this proof is

$$\sum_{x \in B(n)} \Gamma_n(x) = K_n. \tag{4.7}$$

This is valid for the following reason. Let Σ be an open cluster of $B(n)$. Each vertex x of Σ contributes $|\Sigma|^{-1}$ to the sum in (4.7), so that the aggregate contribution from the vertices of Σ equals 1.

We have from (4.6) and (4.7) that

$$\frac{1}{|B(n)|} K_n \ge \frac{1}{|B(n)|} \sum_{x \in B(n)} \Gamma(x). \tag{4.8}$$

However, $(\Gamma(x) : x \in \mathbb{Z}^d)$ is a collection of bounded functions of independent edge-states, and is stationary under translations of the lattice $\mathbb{L}^d$. It follows from an appropriate ergodic theorem (see Dunford and Schwartz (1958, Theorem VIII.6.9) or Tempel'man (1972, Theorem 6.1 and Corollary 6.2)) that

$$\frac{1}{|B(n)|} \sum_{x \in B(n)} \Gamma(x) \to E_p\big(\Gamma(0)\big) \quad \text{a.s.} \tag{4.9}$$

as $n \to \infty$, giving by (4.8) that

$$\liminf_{n\to\infty} \left\{ \frac{1}{|B(n)|} K_n \right\} \geq E_p\big(\Gamma(0)\big) = \kappa(p) \quad \text{a.s.} \tag{4.10}$$

It is not difficult to construct a useful upper bound for K_n. Clearly

$$\begin{aligned} \sum_{x\in B(n)} \Gamma_n(x) &= \sum_{x\in B(n)} \Gamma(x) + \sum_{x\in B(n)} \big\{\Gamma_n(x) - \Gamma(x)\big\} \\ &\leq \sum_{x\in B(n)} \Gamma(x) + \sum_{x:x\leftrightarrow\partial B(n)} \Gamma_n(x) \end{aligned} \tag{4.11}$$

by (4.6), where the final summation is over all vertices x of $B(n)$ which are joined to the surface $\partial B(n) = \{y \in B(n) : \|y\| = n\}$ by open paths. The inequality is valid since $C_n(x) = C(x)$, and therefore $\Gamma_n(x) = \Gamma(x)$, whenever there is no open path joining x to any vertex of $\partial B(n)$. Now, the final sum of (4.11) equals the number of open clusters of $B(n)$ which contain vertices in $\partial B(n)$, and this is no larger than its cardinality $|\partial B(n)|$. Therefore,

$$\begin{aligned} \frac{1}{|B(n)|} \sum_{x\in B(n)} \Gamma_n(x) &\leq \frac{1}{|B(n)|} \left\{ \sum_{x\in B(n)} \Gamma(x) \right\} + \frac{|\partial B(n)|}{|B(n)|} \\ &\to \kappa(p) \quad \text{a.s.} \end{aligned} \tag{4.12}$$

as $n \to \infty$, by (4.9). We combine this with (4.7) and (4.10) to deduce that

$$\frac{1}{|B(n)|} K_n \to \kappa(p) \quad \text{a.s. as } n \to \infty. \tag{4.13}$$

Convergence in L^1 is immediate since $0 \leq K_n |B(n)|^{-1} \leq 1$. □

4.2 Lattice Animals and Large Deviations

Each of the principal characters θ, χ, and κ of percolation theory may be expressed in terms of the distribution of the number $|C|$ of vertices in the open cluster at the origin. To calculate or estimate this distribution is a question involving combinatorics.

We begin by introducing some notation. An *animal* is defined to be a finite connected subgraph of $\mathbb{L}^d$ containing the origin; see Figure 4.1 for two representations of the same two-dimensional animal. If A is an animal, we write A_v and A_e for the sets of vertices and edges of A, respectively, and ΔA for the set of edges of $\mathbb{L}^d$ which do not belong to A but which have at least one endvertex in A; ΔA is called the *edge boundary* of A. We write $\mathcal{A}_{nmb}$ for the set of animals A with

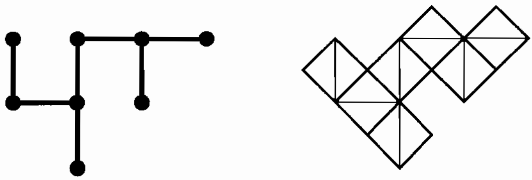


Figure 4.1. The left picture is a two-dimensional animal with eight vertices and seven edges. If we replace each edge by a square having the edge as diagonal, then we obtain a picture which emphasizes the cellular structure of the animal.

$|A_v| = n$, $|A_e| = m$, and $|\Delta A| = b$, and we let $a_{nmb} = |\mathcal{A}_{nmb}|$ be the number of such animals. It is easy to see that $a_{nmb} = 0$ unless n, m, and b satisfy certain inequalities. For example, we may assume that

$$1 \le b \le 2dn \tag{4.14}$$

since each vertex of A is incident to no more than $2d$ edges which are not in A; similarly

$$n - 1 \le m \le dn, \tag{4.15}$$

where the first inequality follows from the fact that every connected graph on n vertices has at least $n - 1$ edges, and the second inequality is a consequence of the observation that each vertex of A has at most $2d$ neighbours in A.

The open cluster C is a random animal:

$$P_p(C = A) = p^m(1-p)^b \quad \text{where } m = |A_e| \text{ and } b = |\Delta A|, \tag{4.16}$$

for all $A \in \mathcal{A}$, the set of all animals. Thus the probability that C contains exactly n vertices is

$$P_p(|C| = n) = \sum_{m,b} a_{nmb} p^m (1-p)^b, \tag{4.17}$$

whence the calculation of the distribution of $|C|$ amounts to finding the generating function of the number of animals. The above series expansion for $P_p(|C| = n)$ is finite, so that $P_p(|C| = n)$ is a finite polynomial in p and is thus analytic. It is not so easy to establish properties of continuity and differentiability for such quantities as the percolation probability $\theta(p)$ or the number of open clusters per vertex,

$$\kappa(p) = \sum_{n=1}^{\infty} \frac{1}{n} P_p(|C| = n), \tag{4.18}$$

and it is with this type of problem in mind that we study the behaviour of a_{nmb} when n, m, and b are large.

Suppose $0 < p = 1-q < 1$ and differentiate (4.17) with respect to p to obtain

$$\frac{d}{dp} P_p(|C| = n) = \sum_{m,b} a_{nmb} \left(\frac{m}{p} - \frac{b}{q} \right) p^m q^b. \tag{4.19}$$

For large n, the behaviour of this sum depends on the typical behaviour of the quantity $p^{-1}|A_e| - q^{-1}|\Delta A|$ as A ranges over the set of animals with n vertices. We shall estimate this quantity with the aid of the following 'large deviation' estimate.

(4.20) Theorem. *There exists $\varepsilon > 0$ such that, for $0 < x < \varepsilon$,*

$$\sum_{m,b:\left|\frac{m}{p} - \frac{b}{q}\right| > dxn} a_{nmb} p^m q^b \le 3d^2 n^2 \exp(-\tfrac{1}{3} n x^2 p^2 q) \tag{4.21}$$

whenever $n \ge 1$ and $0 < p = 1 - q < 1$.

Rewritten in terms of probabilities, Theorem (4.20) asserts that

$$P_p\big(|h(C)| > dxn,\ |C| = n\big) \le 3d^2 n^2 \exp(-\tfrac{1}{3} n x^2 p^2 q) \tag{4.22}$$

for all n and all sufficiently small x, where

$$h(C) = \frac{1}{q} |\Delta C| - \frac{1}{p} |C_e|, \tag{4.23}$$

C_e and ΔC being the edge set and edge boundary of C, respectively. This is tantamount to a large-deviation result, since it concerns the probability that $h(C)$ is comparable with the size of C.

We shall not use probabilistic arguments to demonstrate this result; instead, we emphasize the intrinsically combinatorial aspect of percolation theory by using estimates for the numbers of animals.

It is well known that the number of animals with n vertices grows at most exponentially as $n \to \infty$. More explicitly, we shall see in the proof of Theorem (4.20) that the total number of animals with n vertices satisfies

$$\sum_{m,b} a_{nmb} \le 7^{dn} \qquad \text{for all } n. \tag{4.24}$$

As an application of Theorem (4.20), we shall prove in the next section that the number of open clusters per vertex is differentiable at all values of p, including the critical point $p = p_c$.

Proof of Theorem (4.20). We follow Kesten (1982, p. 85). From (4.17),

$$P_p(|C| < \infty) = \sum_{n,m,b} a_{nmb} p^m q^b \le 1, \tag{4.25}$$

whenever $0 \le p = 1 - q \le 1$. We prove (4.24) first. For each fixed n,

$$\sum_{m,b} a_{nmb} p^m q^b \ge \sum_{m,b} a_{nmb} p^{dn} q^{2dn} \qquad \text{by (4.14)–(4.15)}$$
$$= (pq^2)^{dn} \sum_{m,b} a_{nmb}$$

for all values of p, giving by (4.25) that

$$\sum_{m,b} a_{nmb} \le (pq^2)^{-dn}$$

for all p. We choose p to maximize $pq^2 = p(1-p)^2$, and obtain

$$\sum_{m,b} a_{nmb} \le \left(\tfrac{27}{4}\right)^{dn},$$

which is ample. Now we return to the proof of the theorem.

We have from (4.25) that $a_{nmb} p^m (1-p)^b \le 1$ for all n, m, b, and all p. We may choose $p = m(m+b)^{-1}$, $q = b(m+b)^{-1}$, to obtain

$$(4.26) \qquad a_{nmb} \le p^{-m} q^{-b} = \left(\frac{m+b}{m}\right)^m \left(\frac{m+b}{b}\right)^b \qquad \text{for } n \ge 2.$$

Actually, the right side here is an upper bound for the number $\sum_n a_{nmb}$ of animals with m edges and b boundary edges, but this fact will not be used.

Now, for fixed $n \ge 2$,

(4.27)
$$\sum_{m,b:\left|\frac{m}{p}-\frac{b}{q}\right|>dxn} a_{nmb} p^m q^b$$
$$\le \sum_{m,b:\left|\frac{m}{p}-\frac{b}{q}\right|>dxn} \left(\frac{m+b}{m}\right)^m \left(\frac{m+b}{b}\right)^b p^m q^b \qquad \text{by (4.26)}$$
$$\le 2d^2 n^2 \max\{f(m,b)\}$$

by (4.14) and (4.15), where

$$(4.28) \qquad f(m,b) = \left(\frac{(m+b)p}{m}\right)^m \left(\frac{(m+b)q}{b}\right)^b$$

and the maximum is over all m and b satisfying (4.14), (4.15), and the inequality $|mq - bp| > dxnpq$. From (4.15), this maximum is no larger than the maximum of $f(m,b)$ over m and b satisfying $|mq - bp| \geq xmpq$. We fix m and think of $f(m,y)$ as a function of the continuous variable y. Let

$$\begin{aligned}(4.29)\qquad g_m(y) &= \log f(m,y)\\ &= m\log\left(\frac{(m+y)p}{m}\right) + y\log\left(\frac{(m+y)q}{y}\right).\end{aligned}$$

We differentiate g_m with respect to y to find that

$$g'_m(y) = \log\left(\frac{(m+y)q}{y}\right) = \log\left(1 + \frac{mq - yp}{y}\right),$$

giving that g_m is an increasing function of y if $mq - yp > 0$ and a decreasing function of y if $mq - yp < 0$. Thus the maximum value of $g_m(y)$ over values of y satisfying $|mq - yp| \geq xmpq$ occurs either when $mq - yp = xmpq$ or when $mq - yp = -xmpq$. These two possibilities correspond to

$$(4.30)\qquad y = mqp^{-1}(1 \pm xp).$$

We substitute these two values into (4.29) and expand the logarithms to find that

$$\begin{aligned}g_m(y) &= m\log(1 \pm xpq) + mqp^{-1}(1 \pm xp)\log\left(\frac{1 \pm xpq}{1 \pm xp}\right)\\ &= -\tfrac{1}{2}mx^2p^2q\big(1 + \mathrm{O}(x)\big),\end{aligned}$$

where the $\mathrm{O}(x)$ term does not depend on m or p. We have proved that, for fixed m and n (≥ 2) satisfying $m \leq dn$, it is the case that

$$\max\Big\{f(m,b) : |mp^{-1} - bq^{-1}| > dxn\Big\} \leq \exp(-\tfrac{1}{3}mx^2p^2q)$$

for all small positive values of x. We substitute this into (4.27), and we use (4.15) to obtain

$$\begin{aligned}\sum_{m,b:\left|\frac{m}{p}-\frac{b}{q}\right|>dxn} a_{nmb}p^mq^b &\leq 2d^2n^2\max\Big\{\exp(-\tfrac{1}{3}mx^2p^2q) : m \geq n-1\Big\}\\ &\leq 3d^2n^2\exp(-\tfrac{1}{3}nx^2p^2q) \qquad \text{if } n \geq 2\end{aligned}$$

for all small positive values of x. It is straightforward to handle the term in $n = 1$, and we omit the details of this. □

4.3 Differentiability of κ

As an application of the techniques of the last section concerning the counting of lattice animals, we prove next that κ is a continuously differentiable function of p throughout the interval $[0, 1]$.

(4.31) Theorem. Differentiability of κ. *The number $\kappa(p)$ of open clusters per vertex is a continuously differentiable function on* $[0, 1]$.

We shall prove that the derivative of κ is the term by term derivative of the series in (4.18):

$$\kappa'(p) = \sum_{n,m,b} \frac{1}{n} a_{nmb}\left(mp^{m-1}q^{b} - bp^{m}q^{b-1}\right), \tag{4.32}$$

where $q = 1 - p$.

Rather more than the conclusion of Theorem (4.31) is true. We shall see in Sections 6.4 and 8.7 that κ is analytic on $[0, p_c)$ and infinitely differentiable on $(p_c, 1]$; it is believed that κ is analytic on $(p_c, 1]$ also. Furthermore, it is thought that κ is twice differentiable at p_c but not three times differentiable; this question is related to the problem of critical exponents, to which we return in Chapter 9. Somewhat more is known in two dimensions: by self-duality (see Theorem (11.4)) κ is analytic on $(p_c, 1]$, and in addition Kesten (1982) has shown that κ is twice differentiable at $p_c(2)$ in this case.

Proof of Theorem (4.31). If we differentiate the series in (4.18) term by term we obtain

$$\sum_{n=1}^{\infty} n^{-1} \frac{d}{dp} P_p(|C| = n);$$

we show first that this series converges uniformly for values of p in every interval of the form $[p_1, p_2]$ where $0 < p_1 < p_2 < 1$. Suppose that $0 < p = 1 - q < 1$. Then

$$\frac{d}{dp} P_p(|C| = n) = \sum_{m,b} a_{nmb} \left(\frac{m}{p} - \frac{b}{q}\right) p^m q^b \tag{4.33}$$

from (4.19). We split this sum into two parts depending on whether or not it is the case that $|mp^{-1} - bq^{-1}| > dxn$, where x is small and will be chosen later. When $|mp^{-1} - bq^{-1}| > dxn$, we use the inequality

$$\left|\frac{m}{p} - \frac{b}{q}\right| \le \frac{m+b}{pq} \le \frac{3dn}{pq}, \tag{4.34}$$

valid because of (4.14) and (4.15). By the result of Theorem (4.20), when x is small,

$$\left|n^{-1} \frac{d}{dp} P_p(|C| = n)\right| \le dx P_p(|C| = n) + \left(\frac{3d}{pq}\right) 3d^2n^2 \exp(-\tfrac{1}{3}nx^2p^2q).$$

We set $x = a(n^{-1}\log n)^{1/2}$ to obtain, for $n \geq N \geq 3$,

$$\left| n^{-1}\frac{d}{dp}P_p(|C| = n)\right| \leq da\sqrt{\frac{\log N}{N}}P_p(|C| = n) + \gamma n^{2-a^2\eta}$$

for appropriate functions $\gamma = \gamma(p, d) < \infty$ and $\eta = \eta(p) > 0$ which are continuous in p. Choose p_1 and p_2 satisfying $0 < p_1 < p_2 < 1$, write

$$\widehat{\gamma} = \sup\{\gamma(p, d) : p_1 \leq p \leq p_2\}, \quad \widehat{\eta} = \inf\{\eta(p) : p_1 \leq p \leq p_2\},$$

and choose a such that $2 - a^2\widehat{\eta} < -2$. Then

$$\text{(4.35)} \quad \sum_{n=N}^{\infty}\left| n^{-1}\frac{d}{dp}P_p(|C| = n)\right| \leq da\sqrt{\frac{\log N}{N}} + \widehat{\gamma}\sum_{n=N}^{\infty}\frac{1}{n^2} \quad \text{if } p_1 \leq p \leq p_2,$$

since

$$\sum_{n=N}^{\infty} P_p(|C| = n) \leq 1.$$

It follows that the left side of (4.35) converges to 0 as $N \to \infty$, uniformly in $p \in [p_1, p_2]$, and therefore κ is continuously differentiable on $(0, 1)$ with derivative given by (4.32).

It remains to show that κ is continuously differentiable at $p = 0$ and $p = 1$. We shall not spend much time on this small point, since it is not particularly important and furthermore we shall prove much more later (see Theorems (6.108) and (8.92)). Here is a sketch of a primitive argument for the case $p = 0$; an analogous argument is valid for $p = 1$. We have that $P_p(|C| < \infty) = 1$ when $p < p_c$, and therefore,

$$\sum_{n,m,b} a_{nmb}p^m q^b = 1 \qquad \text{for } p < p_c.$$

Clearly $\kappa(0) = 1$, so that

$$\frac{\kappa(p) - \kappa(0)}{p} = \sum_{n,m,b:n\geq 2}\left(\frac{1}{n} - 1\right)a_{nmb}p^{m-1}q^b.$$

We pick out the term in p^0:

$$\text{(4.36)} \qquad \frac{\kappa(p) - \kappa(0)}{p} = A(p) + \sum_{n,m,b:m=1}\left(\frac{1}{n} - 1\right)a_{nmb}p^{m-1}q^b,$$

where

$$A(p) = \sum_{n,m,b:m\geq 2}\left(\frac{1}{n} - 1\right)a_{nmb}p^{m-1}q^b$$
$$\to 0 \qquad \text{as } p \to 0;$$

this is not difficult to check (use (4.24)). There are exactly $2d$ animals containing the origin which have $m = 1$, and each such animal has $n = 2$ and $b = 2(2d-1)$. We let $p \to 0$ in (4.36) to find that $\kappa'(0) = -d$. On the other hand, we have from the first part of the proof that

$$(4.37) \qquad \kappa'(p) = \sum_{n,m,b} \frac{1}{n} a_{nmb}\big(mp^{m-1}q^b - bp^m q^{b-1}\big) \qquad \text{if } 0 < p < 1.$$

We may use an argument similar to the above to see that $\kappa'(p) \to -d$ as $p \to 0$, and we omit the details of this. □

4.4 Notes

Section 4.1. Sykes and Essam (1964) used the number of open clusters per vertex in their beautiful but non-rigorous derivation of exact values for the critical probabilities of certain two-dimensional lattices (see Section 11.2). The existence of the limit of Theorem (4.2) was proved by Grimmett (1976), and the form of the limit by Wierman (1978); the proof here is essentially that of Grimmett (1981b) with an improvement due to Kesten (1982, p. 240). Brånvall (1980), Cox and Grimmett (1981, 1984), and Zhang (1998) have proved central limit theorems for such quantities as the number of open clusters within a large box. See Section 11.6 for a discussion of the central limit theorem when $d = 2$.

Bezuidenhout, Grimmett, and Löffler (1998) have studied the mean number of clusters per vertex for spread out percolation models and continuum percolation.

Section 4.2. The large deviation result of Theorem (4.20) is due to Kesten (1982, p. 85). Aizenman, Kesten, and Newman (1987a) attain a similar goal by related methods.

We mention the beautiful theorem of Delyon (1980). In a general context, he explores the asymptotic behaviour of the number of animals with m edges and αm boundary edges, showing that this number grows at a rate which depends on neither the dimension of the graph nor the lattice structure, so long as α is no larger than some value *which depends* on the choice of lattice.

Section 4.3. Grimmett (1981b) first used lattice animals to show that κ is differentiable in two dimensions; Kesten (1982) carried such calculations further. See Aizenman, Kesten, and Newman (1987a) and Nguyen (1988) also. Yang and Zhang (1992) have explored the differentiability of κ near the critical point for large d.

It is an open problem to show that κ is not infinitely differentiable at p_c, for any $d \geq 2$.

Chapter 5

Exponential Decay

5.1 Mean Cluster Size

Let us consider bond percolation on $\mathbb{L}^d$ with edge-density p, where $d \geq 2$. As we have remarked, the following two functions are two of the principal characters in the action:

$$\theta(p) = P_p(|C| = \infty), \qquad \chi(p) = E_p|C|,$$

where C is the open cluster containing the origin, and $|C|$ is the number of vertices in C. We have seen that there exists $p_c = p_c(d)$ in $(0, 1)$ such that

$$\theta(p)\begin{cases} = 0 & \text{if } p < p_c, \\ > 0 & \text{if } p > p_c. \end{cases} \tag{5.1}$$

The critical value p_c is the value of p above which infinite open clusters appear. Much of the past romance of percolation theory has been associated with the question of whether or not p_c is the unique value of p at which a critical phenomenon occurs; in particular, we shall see in Chapter 11 that relatively simple arguments imply that $p_c(2) = \frac{1}{2}$ once we have answered this question affirmatively.

The function χ experiences a critical phenomenon also. Clearly χ is non-decreasing in p, and $\chi(0) = 1$, $\chi(1) = \infty$. We have, as in Section 1.5, that

$$\chi(p) = \infty \cdot P_p(|C| = \infty) + \sum_{n=1}^{\infty} n P_p(|C| = n),$$

so that $\chi(p) = \infty$ whenever $\theta(p) > 0$. Thus $\chi(p) = \infty$ when $p > p_c$. It is not so easy to show that $\chi(p) < \infty$ when $p < p_c$, and this result is one of the principal targets of this chapter.

(5.2) Theorem. *The mean size $\chi(p)$ of the open cluster containing the origin is finite if $p < p_c$.*

One of the greatest challenges in percolation theory has been to prove this result, which has in the past been reformulated as follows. We define a second 'critical probability' $p_T = p_T(d)$ by

$$p_T = \sup\{p : \chi(p) < \infty\}, \tag{5.3}$$

the value of p at which χ changes from being finite to being infinite. It follows from the remarks above that $p_T \le p_c$, and the problem was to prove that $p_T = p_c$. The special case of this conjecture for the planar lattice $\mathbb{L}^2$ is of particular historical interest, since this was the major obstacle in the way of a complete proof that the critical probability $p_c(2)$ equals $\frac{1}{2}$. With this exact calculation as a target, the conjecture attracted many individuals by its simplicity but apparent difficulty. It is relatively easy to show that $p_T(2) + p_c(2) \le 1$, and it has been known since 1960 that $p_c(2) \ge \frac{1}{2}$, so that it remained to show the equality of $p_T(2)$ and $p_c(2)$; there is therefore a sense in which Theorem (5.2) is the natural generalization to higher dimensions of the exact calculation in two dimensions.

We present two results each of which implies Theorem (5.2). The first is due to Menshikov (1986) (see also Menshikov, Molchanov, and Sidorenko (1986)), and is presented in Section 5.2. Menshikov has proved the stronger result that the tail of the radius of the open cluster at the origin decays exponentially whenever $p < p_c$. As a by-product of the proof, we obtain that there exists a (> 0) such that $\theta(p) - \theta(p_c) \ge a(p - p_c)$ when $p - p_c$ is small and positive; this observation will be of value in the discussion concerning critical exponents (see Chapters 9 and 10). The second proof of Theorem (5.2) is due to Aizenman and Barsky (1987) and is perhaps slightly less basic and intuitive than the first proof; it is presented in Section 5.3. At first sight these two proofs may appear to differ fundamentally, but actually they have a lot in common. The Aizenman–Barsky proof yields also the fact that θ grows at least linearly in $p - p_c$, as well as an estimate for the decay rate of $P_{p_c}(|C| = n)$ as $n \to \infty$.

5.2 Exponential Decay of the Radius Distribution beneath p_c

Let $S(n)$ be the ball of radius n with centre at the origin: $S(n)$ is the set of all vertices x in $\mathbb{Z}^d$ for which $\delta(0, x) \le n$. We write $\partial S(n)$ for the surface of $S(n)$, being the set of all x with $\delta(0, x) = n$. Let A_n be the event that there exists an open path joining the origin to some vertex in $\partial S(n)$. The following is the main result.

(5.4) Theorem. Exponential tail decay of the radius of an open cluster. *If $p < p_c$, there exists $\psi(p) > 0$ such that*

$$P_p(A_n) < e^{-n\psi(p)} \quad \textit{for all } n. \tag{5.5}$$

This theorem amounts to the assertion that the radius of the cluster at the origin has a tail which decays at least exponentially fast when $p < p_c$. Theorem (5.2) follows immediately, using the following elementary argument. The number of vertices in $S(n)$ is no larger than the volume of the Euclidean ball in $\mathbb{R}^d$ with radius $n + 1$, and so there exists $\nu = \nu(d)$ such that

$$(5.6) \qquad |S(n)| \le \nu(n+1)^d.$$

Let $M = \max\{n : A_n \text{ occurs}\}$. If $p < p_c$ then $P_p(M < \infty) = 1$, giving that

$$\begin{aligned} E_p|C| &\le \sum_n E_p\big(|C| \,\big|\, M = n\big) P_p(M = n) \\ &\le \sum_n |S(n)| P_p(A_n) \le \sum_n \nu(n+1)^d e^{-n\psi(p)} \\ &< \infty, \end{aligned}$$

where $\psi(p)$ is the quantity given in Theorem (5.4).

The above theorem is a lot stronger than is needed to demonstrate Theorem (5.2). It implies not only that $E_p|C| < \infty$ when $p < p_c$, but also an estimate on the tail probabilities of $|C|$. To see this, suppose that $p < p_c$ and $|C| > n$. If $C \subseteq S(m)$ then $n \le \nu(m+1)^d$ by (5.6), giving that $m \ge (n/\nu)^{1/d} - 1$. Thus, if $m < (n/\nu)^{1/d} - 1$ then C is not contained in $S(m)$, giving that C contains some vertex in $\partial S(k)$ where $k = \lfloor (n/\nu)^{1/d} \rfloor$. It follows that

$$(5.7) \qquad \begin{aligned} P_p(|C| > n) &\le P_p(A_k) \\ &\le \exp\big(-n^{1/d}\beta(p)\big) \qquad \text{when } p < p_c, \end{aligned}$$

for some $\beta(p) > 0$. We shall see in Chapter 6 that this estimate is not best possible: the term $n^{1/d}$ in the exponent may be replaced by n.

The principal step in proving Theorem (5.4) involves an estimate which is valid for all $p \in (0, 1)$, and not merely for values of p less than p_c. This estimate has an interesting consequence when $p > p_c$, in that it implies that $\theta(p) - \theta(p_c)$ grows at least linearly in $p - p_c$. We note this fact here, and shall make use of it in Chapter 10 when discussing the matter of 'critical exponents'.

(5.8) Theorem. *There exist $a, b > 0$ such that*

$$\theta(p) - \theta(p_c) \ge a(p - p_c) \qquad \text{if } 0 \le p - p_c \le b.$$

We move on to the proof of Theorem (5.4), which is based on the work of Menshikov (1986). Let $S(n, x)$ be the ball of radius n with centre at the vertex x, and let $\partial S(n, x)$ be the surface of $S(n, x)$; thus $S(n, x) = x + S(n)$ and $\partial S(n, x) = x + \partial S(n)$. Similarly, let $A_n(x)$ be the event that there exists an open path from the vertex x to some vertex in $\partial S(n,x)$. We are concerned with the probabilities

$$g_p(n) = P_p(A_n) = P_p\big(A_n(x)\big) \qquad \text{for any } x.$$

Now A_n is an increasing event which depends on the edges joining vertices in $S(n)$ only. We apply Russo's formula to $P_p(A_n)$ to obtain

$$g_p'(n) = E_p\big(N(A_n)\big) \tag{5.9}$$

where the prime denotes differentiation with respect to p, and $N(A_n)$ is the number of edges which are pivotal for A_n. It follows as in (2.29) that

$$g_p'(n) = \frac{1}{p} E_p\big(N(A_n); A_n\big) = \frac{1}{p} E_p\big(N(A_n) \mid A_n\big) g_p(n)$$

so that

$$\frac{1}{g_p(n)} g_p'(n) = \frac{1}{p} E_p\big(N(A_n) \mid A_n\big). \tag{5.10}$$

Let $0 \le \alpha < \beta \le 1$, and integrate (5.10) from $p = \alpha$ to $p = \beta$ to obtain

$$\begin{aligned} g_\alpha(n) &= g_\beta(n) \exp\left(-\int_\alpha^\beta \frac{1}{p} E_p\big(N(A_n) \mid A_n\big)\, dp\right) \\ &\le g_\beta(n) \exp\left(-\int_\alpha^\beta E_p\big(N(A_n) \mid A_n\big)\, dp\right), \end{aligned} \tag{5.11}$$

as in (2.30). We need now to show that $E_p(N(A_n) \mid A_n)$ grows roughly linearly in n when $p < p_c$, and this inequality will then yield an upper bound for $g_\alpha(n)$ of the form required in (5.5). The vast majority of the work in the proof is devoted to estimating $E_p(N(A_n) \mid A_n)$, and the rationale is roughly as follows. If $p < p_c$ then $P_p(A_n) \to 0$ as $n \to \infty$, so that for large n we are conditioning on an event of small probability. If A_n occurs, 'but only just', then the connections between the origin and $\partial S(n)$ must be sparse; indeed, there must exist many open edges in $S(n)$ which are crucial for the occurrence of A_n (see Figure 5.1). It is plausible that the number of such pivotal edges in paths from the origin to $\partial S(2n)$ is approximately twice the number of such edges in paths to $\partial S(n)$, since these sparse paths have to traverse twice the distance. Thus the number $N(A_n)$ of edges pivotal for A_n should grow linearly in n.

Suppose that the event A_n occurs, and denote by $e_1, e_2, \dots, e_N$ the (random) edges which are pivotal for A_n. Since A_n is increasing, each e_j has the property that A_n occurs if and only if e_j is open; thus all open paths from the origin to $\partial S(n)$ traverse e_j, for every j (see Figure 5.1). Let π be such an open path; we assume that the edges $e_1, e_2, \dots, e_N$ have been enumerated in the order in which they are traversed by π. A glance at Figure 5.1 confirms that this ordering is independent

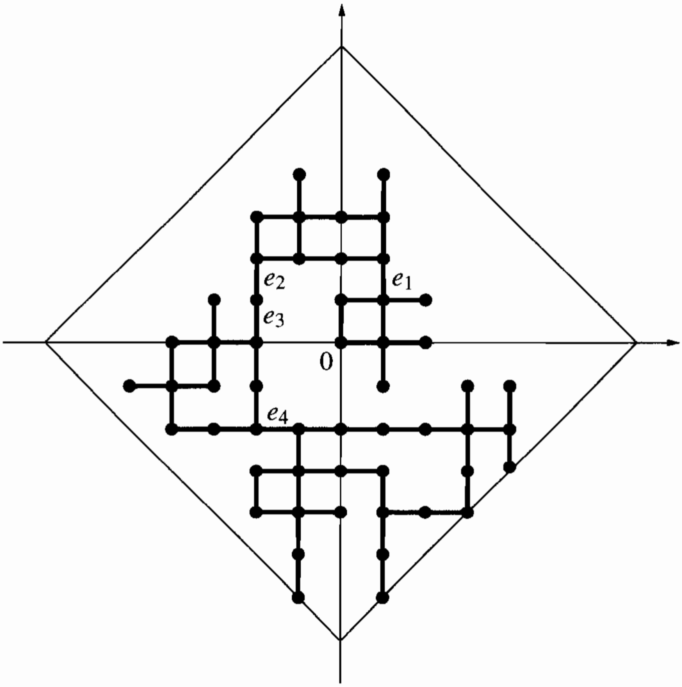


Figure 5.1. A picture of the open cluster of $S(7)$ at the origin. There are exactly four pivotal edges for A_n in this configuration, and these are labelled e_1, e_2, e_3, e_4.

of the choice of π. We denote by x_i the endvertex of e_i encountered first by π, and by y_i the other endvertex of e_i. We observe that there exist at least two edge-disjoint open paths joining 0 to x_1, since, if two such paths cannot be found then, by Menger's theorem (Wilson (1979, p. 126)), there exists a pivotal edge in π which is encountered prior to x_1, a contradiction. Similarly, for $1 \leq i < N$, there exist at least two edge-disjoint open paths joining y_i to x_{i+1}; see Figure 5.2. In the words of the discoverer of this proof, the open cluster containing the origin resembles a chain of sausages.

As before, let $M = \max\{k : A_k \text{ occurs}\}$ be the radius of the largest ball whose surface contains a vertex which is joined to the origin by an open path. We note that, if $p < p_c$, then M has a non-defective distribution in that $P_p(M \geq k) = g_p(k) \to 0$ as $k \to \infty$. We shall show that, conditional on A_n, the number $N(A_n)$ is at least as large as the number of renewals up to time n of a certain renewal process whose inter-renewal times have approximately the same distribution as M. In order to compare $N(A_n)$ with such a renewal process, we introduce the following notation. Let $\rho_1 = \delta(0, x_1)$ and $\rho_{i+1} = \delta(y_i, x_{i+1})$ for $1 \leq i < N$. The first step is to show that, roughly speaking, the random variables $\rho_1, \rho_2, \ldots$ are jointly smaller in distribution than a sequence $M_1, M_2, \ldots$ of independent random variables distributed as M.

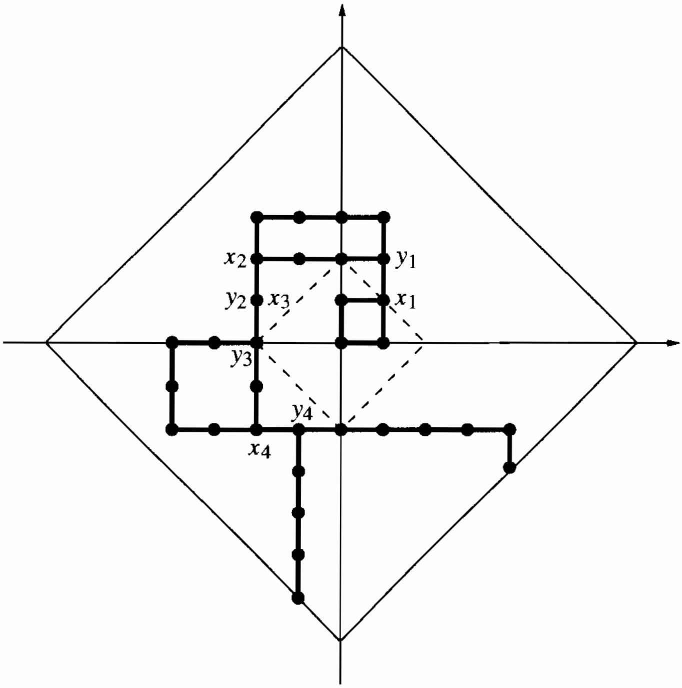


Figure 5.2. The pivotal edges are $e_i = \langle x_i, y_i \rangle$ for $i = 1, 2, 3, 4$. We have that $x_3 = y_2$ in this configuration. The dashed line is the surface $\partial S(\rho_1)$ of $S(\rho_1)$. Note the two edge-disjoint paths from the origin to $\partial S(\rho_1)$.

(5.12) Lemma. *Let k be a positive integer, and let $r_1, r_2, \dots, r_k$ be non-negative integers such that $\sum_{i=1}^{k} r_i \le n - k$. For $0 < p < 1$,*

$$(5.13)\quad P_p\big(\rho_k \le r_k,\ \rho_i = r_i \text{ for } 1 \le i < k \,\big|\, A_n\big) \ge P_p(M \le r_k) P_p\big(\rho_i = r_i \text{ for } 1 \le i < k \,\big|\, A_n\big).$$

Proof. Suppose by way of illustration that $k = 1$ and $0 \le r_1 < n$. Then

$$(5.14)\qquad \{\rho_1 > r_1\} \cap A_n \subseteq A_{r_1+1} \circ A_n,$$

since if $\rho_1 > r_1$ then the first endvertex of the first pivotal edge lies either outside $S(r_1 + 1)$ or on its surface $\partial S(r_1 + 1)$; see Figure 5.2. However, A_{r_1+1} and A_n are increasing events which depend on the edges within $S(n)$ only, and the BK inequality yields

$$P_p\big(\{\rho_1 > r_1\} \cap A_n\big) \le P_p(A_{r_1+1}) P_p(A_n).$$

We divide by $P_p(A_n)$ to obtain

$$P_p\big(\rho_1 > r_1 \,\big|\, A_n\big) \le g_p(r_1 + 1);$$

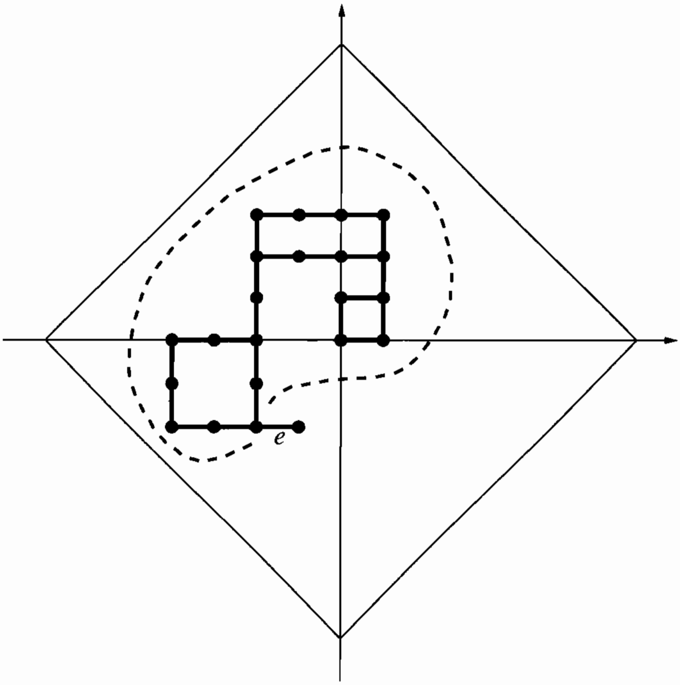


Figure 5.3. A sketch of the event B_e. The dashed curve indicates that the only open 'exit' from its interior is via the edge e. Note the existence of 3 pivotal open edges for the event that 0 is connected to an endvertex of e.

however, $P_p(M \geq m) = g_p(m)$, and thus we have obtained (5.13) in the case $k = 1$.

We now prove the lemma for general values of k. Suppose that $k \geq 1$, and let $r_1, r_2, \ldots, r_k$ be non-negative integers with sum not exceeding $n - k$. Let N be the number of edges which are pivotal for A_n; we enumerate and label these edges as $e_i = \langle x_i, y_i \rangle$ as before. For any edge $e = \langle u, v \rangle$, let D_e be the set of vertices attainable from 0 along open paths not using e, together with all open edges between such vertices. Let B_e be the event that the following statements hold:

(a) exactly one of u or v lies in D_e, say u;

(b) e is open;

(c) D_e contains no vertex of $\partial S(n)$;

(d) the pivotal edges for the event $\{0 \leftrightarrow v\}$ are, taken in order, $\langle x_1, y_1 \rangle$, $\langle x_2, y_2 \rangle$, $\ldots$, $\langle x_{k-2}, y_{k-2} \rangle$, $\langle x_{k-1}, y_{k-1} \rangle = e$, where $\delta(y_{i-1}, x_i) = r_i$ for $1 \leq i < k$, and $y_0 = 0$.

We now define the event $B = \bigcup_e B_e$. For $\omega \in A_n \cap B$, there exists a unique edge $e = e(\omega)$ such that B_e occurs. See Figure 5.3.

For $\omega \in B$, we consider the set of vertices and open edges attainable along open paths from the origin without using $e = e(\omega)$; to this graph we append e and its other endvertex $v = y_{k-1}$, and we place a mark over y_{k-1} in order to distinguish it from the other vertices. We denote by G the resulting (marked) graph, and we write $y(G)$ for the unique marked vertex of G. We condition on G to obtain

$$P_p(A_n \cap B) = \sum_{\Gamma} P_p(B,\ G = \Gamma) P_p(A_n \mid B,\ G = \Gamma),$$

where the sum is over all possible values Γ of G. The final term in this summation is the (conditional) probability that $y(\Gamma)$ is joined to $\partial S(n)$ by an open path which has no vertex other than $y(\Gamma)$ in common with Γ. Thus, in the obvious terminology,

$$(5.15) \qquad P_p(A_n \cap B) = \sum_{\Gamma} P_p(B,\ G = \Gamma) P_p\big(y(\Gamma) \leftrightarrow \partial S(n) \text{ off } \Gamma\big).$$

Similarly,

$$\begin{aligned} &P_p\big(\{\rho_k > r_k\} \cap A_n \cap B\big) \\ &\quad = \sum_{\Gamma} P_p(B,\ G = \Gamma) P_p\Big(\{\rho_k > r_k\} \cap A_n \,\Big|\, B,\ G = \Gamma\Big) \\ &\quad = \sum_{\Gamma} P_p(B,\ G = \Gamma) \\ &\qquad \times P_p\Big(\big\{y(\Gamma) \leftrightarrow \partial S(r_k + 1, y(\Gamma)) \text{ off } \Gamma\big\} \circ \big\{y(\Gamma) \leftrightarrow \partial S(n) \text{ off } \Gamma\big\}\Big). \end{aligned}$$

We apply the BK inequality to the last term to obtain

$$\begin{aligned} (5.16) \quad & P_p\big(\{\rho_k > r_k\} \cap A_n \cap B\big) \\ & \le \sum_{\Gamma} P_p(B,\ G = \Gamma)\, P_p\big(y(\Gamma) \leftrightarrow \partial S(n) \text{ off } \Gamma\big) \\ & \qquad\qquad \times P_p\big(y(\Gamma) \leftrightarrow \partial S(r_k + 1, y(\Gamma)) \text{ off } \Gamma\big) \\ & \le g_p(r_k + 1) P_p(A_n \cap B) \end{aligned}$$

by (5.15) and the fact that, for each possible Γ,

$$\begin{aligned} P_p\big(y(\Gamma) \leftrightarrow \partial S(r_k + 1, y(\Gamma)) \text{ off } \Gamma\big) &\le P_p\big(y(\Gamma) \leftrightarrow \partial S(r_k + 1, y(\Gamma))\big) \\ &= P_p(A_{r_k+1}) \\ &= g_p(r_k + 1). \end{aligned}$$

We divide each side of (5.16) by $P_p(A_n \cap B)$ to obtain

$$P_p\big(\rho_k \le r_k \,\big|\, A_n \cap B\big) \ge 1 - g_p(r_k + 1),$$

throughout which we multiply by $P_p(B \mid A_n)$ to obtain the result. □

(5.17) Lemma. *For $0 < p < 1$, it is the case that*

(5.18) $$E_p\big(N(A_n)\,\big|\,A_n\big) \geq \frac{n}{\sum_{i=0}^{n} g_p(i)} - 1.$$

Proof. It follows from Lemma (5.12) that

(5.19) $$P_p\big(\rho_1+\rho_2+\cdots+\rho_k \leq n-k\,\big|\,A_n\big) \geq P(M_1+M_2+\cdots+M_k \leq n-k),$$

where $k \geq 1$ and $M_1, M_2, \ldots$ is a sequence of independent random variables distributed as M. We defer until the end of this proof the minor chore of deducing (5.19) from (5.13). Now $N(A_n) \geq k$ if $\rho_1 + \rho_2 + \cdots + \rho_k \leq n - k$, so that

(5.20) $$P_p\big(N(A_n) \geq k\,\big|\,A_n\big) \geq P(M_1 + M_2 + \cdots + M_k \leq n - k).$$

A minor difficulty is that the M_i may have a defective distribution. Indeed,

$$P(M \geq r) = g_p(r) \to \theta(p) \qquad \text{as } r \to \infty;$$

thus we allow the M_i to take the value ∞ with probability $\theta(p)$. On the other hand, we are not concerned with atoms at ∞, since

$$P(M_1 + M_2 + \cdots + M_k \leq n - k) = P(M_1' + M_2' + \cdots + M_k' \leq n),$$

where $M_i' = 1+\min\{M_i, n\}$, and we work henceforth with these truncated random variables. Summing (5.20) over k, we obtain

(5.21) $$\begin{aligned} E_p\big(N(A_n)\,\big|\,A_n\big) &\geq \sum_{k=1}^{\infty} P(M_1' + M_2' + \cdots + M_k' \leq n) \\ &= \sum_{k=1}^{\infty} P(K \geq k+1) \\ &= E(K) - 1, \end{aligned}$$

where

$$K = \min\{k : M_1' + M_2' + \cdots + M_k' > n\}.$$

Let $S_k = M_1' + M_2' + \cdots + M_k'$, the sum of independent, identically distributed, bounded random variables. By Wald's equation (see Chow and Teicher (1978, pp. 137, 150) or Grimmett and Stirzaker (1992, pp. 396, 466)),

$$n < E(S_K) = E(K)E(M_1'),$$

giving that

$$E(K) > \frac{n}{E(M_1')} = \frac{n}{1 + E(\min\{M_1, n\})} = \frac{n}{\sum_{i=0}^{n} g_p(i)}$$

since

$$E\big(\min\{M_1, n\}\big) = \sum_{i=1}^{n} P(M \geq i) = \sum_{i=1}^{n} g_p(i).$$

It remains to show that (5.19) follows from Lemma (5.12). We have that

$$\begin{aligned} &P_p\big(\rho_1 + \rho_2 + \cdots + \rho_k \leq n - k \,\big|\, A_n\big) \\ &\quad = \sum_{i=0}^{n-k} P_p\big(\rho_1 + \rho_2 + \cdots + \rho_{k-1} = i,\ \rho_k \leq n - k - i \,\big|\, A_n\big) \\ &\quad \geq \sum_{i=0}^{n-k} P(M \leq n - k - i) P_p\big(\rho_1 + \rho_2 + \cdots + \rho_{k-1} = i \,\big|\, A_n\big) \quad \text{by (5.13)} \\ &\quad = P_p\big(\rho_1 + \rho_2 + \cdots + \rho_{k-1} + M_k \leq n - k \,\big|\, A_n\big), \end{aligned}$$

where M_k is a random variable which is independent of all edge-states in $S(n)$ and is distributed as M. There is a mild abuse of notation here, since P_p is not the correct probability measure unless M_k is measurable on the usual σ-field of events, but we need not trouble ourselves overmuch about this. We iterate the above argument in the obvious way to deduce (5.19), thereby completing the proof of the lemma. □

The conclusion of Theorem (5.8) is easily obtained from this lemma, but we delay this step until the end of the section. The proof of Theorem (5.4) proceeds by substituting (5.18) into (5.11) to obtain that, for $0 \leq \alpha < \beta \leq 1$,

$$g_\alpha(n) \leq g_\beta(n) \exp\left(- \int_\alpha^\beta \left[\frac{n}{\sum_{i=0}^{n} g_p(i)} - 1 \right] dp \right).$$

It is difficult to calculate the integral in the exponent, and so we use the inequality $g_p(i) \leq g_\beta(i)$ for $p \leq \beta$ to obtain

$$g_\alpha(n) \leq g_\beta(n) \exp\left(-(\beta - \alpha) \left[\frac{n}{\sum_{i=0}^{n} g_\beta(i)} - 1 \right] \right), \tag{5.22}$$

and it is from this relation that the conclusion of Theorem (5.4) will be extracted. Before continuing, it is interesting to observe that, by combining (5.10) and (5.18), we obtain a differential–difference inequality involving the function

$$G(p, n) = \sum_{i=0}^{n} g_p(i);$$

rewriting this equation rather informally as a partial differential inequality, we obtain

$$\frac{\partial^2 G}{\partial p\, \partial n} \geq \frac{\partial G}{\partial n} \left(\frac{n}{G} - 1 \right). \tag{5.23}$$

Efforts to integrate this inequality directly and simply have failed so far.

Once we know that

$$E_\beta(M) = \sum_{i=1}^{\infty} g_\beta(i) < \infty \qquad \text{for all } \beta < p_c,$$

then (5.22) gives us that

$$g_\alpha(n) \le e^{-n\psi(\alpha)} \qquad \text{for all } \alpha < p_c,$$

for some $\psi(\alpha) > 0$, as required. At this stage of the proof we know rather less than the finite summability of the $g_p(i)$ for $p < p_c$, knowing only that $g_p(i) \to 0$ as $i \to \infty$. In order to estimate the rate at which $g_p(i) \to 0$, we shall use (5.22) as a mathematical turbocharger.

(5.24) Lemma. *For $p < p_c$, there exists $\delta(p)$ such that*

(5.25) $$g_p(n) \le \delta(p) n^{-1/2} \qquad \textit{for } n \ge 1.$$

Once this lemma has been proved, the theorem follows quickly. To see this, note that (5.25) implies the existence of $\Delta(p) < \infty$ such that

(5.26) $$\sum_{i=0}^{n} g_p(i) \le \Delta(p) n^{1/2} \qquad \text{for } p < p_c.$$

Let $\alpha < p_c$, and find β such that $\alpha < \beta < p_c$. Substitute (5.26) with $p = \beta$ into (5.22) to find that

$$\begin{aligned} g_\alpha(n) &\le g_\beta(n) \exp\left\{-(\beta - \alpha)\left(\frac{n^{1/2}}{\Delta(\beta)} - 1\right)\right\} \\ &\le \exp\left\{1 - \frac{(\beta - \alpha)}{\Delta(\beta)} n^{1/2}\right\}. \end{aligned}$$

Therefore,

$$\sum_{n=1}^{\infty} g_\alpha(n) < \infty \qquad \text{for } \alpha < p_c,$$

and the theorem follows from the observations made prior to the statement of Lemma (5.24). We shall now prove this lemma.

Proof. First, we shall show the existence of a subsequence $n_1, n_2, \ldots$ along which $g_p(n)$ approaches 0 rather quickly; secondly, we shall fill in the gaps in this subsequence.

Fix $\beta < p_c$ and a positive integer n. Let α satisfy $0 < \alpha < \beta$ and let $n' \geq n$; later we shall choose α and n' explicitly in terms of β and n. From (5.22),

$$(5.27) \qquad \begin{aligned} g_\alpha(n') &\leq g_\beta(n') \exp\left(1 - \frac{n'(\beta - \alpha)}{\sum_{i=0}^{n'} g_\beta(i)}\right) \\ &\leq g_\beta(n) \exp\left(1 - \frac{n'(\beta - \alpha)}{\sum_{i=0}^{n'} g_\beta(i)}\right) \end{aligned}$$

since $n \leq n'$. We wish to write the exponent in terms of $g_\beta(n)$, and to this end we shall choose n' appropriately. We split the summation into two parts corresponding to $i < n$ and $i \geq n$, and we use the monotonicity of $g_\beta(i)$ to find that

$$\begin{aligned} \frac{1}{n'} \sum_{i=0}^{n'} g_\beta(i) &\leq \frac{1}{n'} \left\{ n g_\beta(0) + n' g_\beta(n) \right\} \\ &\leq 3 g_\beta(n) \qquad \text{if } n' \geq n \lfloor g_\beta(n)^{-1} \rfloor. \end{aligned}$$

We now define

$$(5.28) \qquad n' = n \gamma_\beta(n) \qquad \text{where } \gamma_\beta(n) = \lfloor g_\beta(n)^{-1} \rfloor$$

and deduce from (5.27) that

$$(5.29) \qquad g_\alpha(n') \leq g_\beta(n) \exp\left(1 - \frac{\beta - \alpha}{3 g_\beta(n)}\right).$$

Next we choose α by setting

$$(5.30) \qquad \beta - \alpha = 3 g_\beta(n) \left\{ 1 - \log g_\beta(n) \right\}.$$

Now $g_\beta(m) \to 0$ as $m \to \infty$, so that $0 < \alpha < \beta$ if n has been picked sufficiently large; (5.29) then yields

$$(5.31) \qquad g_\alpha(n') \leq g_\beta(n)^2.$$

This conclusion is the basic recursion step which we shall use repeatedly. We have shown that, for $\beta < p_c$, there exists $n_0(\beta)$ such that (5.31) holds for all $n \geq n_0(\beta)$ whenever n' and α are given by (5.28) and (5.30), respectively.

Next, we fix $p < p_c$ and choose π such that $p < \pi < p_c$. We now construct sequences $(p_i : i \geq 0)$ of probabilities and $(n_i : i \geq 0)$ of integers as follows. We set $p_0 = \pi$ and shall pick n_0 later. Having found $p_0, p_1, \ldots, p_i$ and $n_0, n_1, \ldots, n_i$, we define

$$(5.32) \qquad n_{i+1} = n_i \gamma_i \quad \text{and} \quad p_i - p_{i+1} = 3 g_i (1 - \log g_i)$$

where $g_i = g_{p_i}(n_i)$ and $\gamma_i = \lfloor g_i^{-1} \rfloor$. We note that $n_i \le n_{i+1}$ and $p_i > p_{i+1}$. The recursion (5.32) is valid so long as $p_{i+1} > 0$, and this is indeed the case so long as n_0 has been chosen to be sufficiently large. To see this we argue as follows. From the definition of $p_0, \dots, p_i$ and $n_0, \dots, n_i$ and the discussion leading to (5.31), we find that

$$g_{j+1} \le g_j^2 \quad \text{for } j = 0, 1, \dots, i-1. \tag{5.33}$$

If a real sequence $(x_j : j \ge 0)$ satisfies $0 < x_0 < 1$ and $x_{j+1} = x_j^2$ for $j \ge 0$, it is easy to check that

$$s(x_0) = \sum_{j=0}^{\infty} 3x_j(1 - \log x_j) < \infty,$$

and furthermore that $s(x_0) \to 0$ as $x_0 \to 0$. We may pick x_0 sufficiently small such that

$$s(x_0) \le \pi - p \tag{5.34}$$

and then we pick n_0 sufficiently large that $g_0 = g_\pi(n_0) < x_0$. Now $h(x) = 3x(1 - \log x)$ is an increasing function on $[0, x_0]$, giving from (5.32) and (5.33) that

$$\begin{aligned} p_{i+1} &= p_i - 3g_i(1 - \log g_i) \\ &= \pi - \sum_{j=0}^{i} 3g_j(1 - \log g_j) \\ &\ge \pi - \sum_{j=0}^{\infty} 3x_j(1 - \log x_j) \\ &\ge p \qquad \text{by (5.34).} \end{aligned}$$

Thus, by a suitable choice of n_0, we may guarantee not only that $p_{i+1} > 0$ for all i but also that

$$\widetilde{p} = \lim_{i \to \infty} p_i$$

satisfies $\widetilde{p} \ge p$. Let us suppose that n_0 has been chosen accordingly, so that the recursion (5.32) is valid and $\widetilde{p} \ge p$. We have from (5.32) and (5.33) that

$$n_k = n_0 \gamma_0 \gamma_1 \dots \gamma_{k-1} \qquad \text{for } k \ge 1$$

and

$$\begin{aligned} g_{k-1}^2 &= g_{k-1} g_{k-1} \le g_{k-1} g_{k-2}^2 \le \cdots \\ &\le g_{k-1} g_{k-2} \cdots g_1 g_0^2 \\ &\le (\gamma_{k-1}\gamma_{k-2} \cdots \gamma_0)^{-1} g_0 \\ &= \delta^2 n_k^{-1}, \end{aligned} \tag{5.35}$$

where $\delta^2 = n_0 g_0$.

We are essentially finished. Let $n > n_0$, and find an integer k such that $n_{k-1} \le n < n_k$; this is always possible since $g_k \to 0$ as $k \to \infty$, and therefore $n_{k-1} < n_k$ for all large k. Then

$$\begin{aligned} g_p(n) &\le g_{p_{k-1}}(n_{k-1}) && \text{since } p \le p_{k-1} \\ &= g_{k-1} \\ &\le \delta n_k^{-1/2} && \text{by (5.35)} \\ &\le \delta n^{-1/2} && \text{since } n < n_k \end{aligned}$$

as required. This is valid for $n > n_0$, but we may adjust the constant δ so that a similar inequality is valid for all $n \ge 1$. □

Proof of Theorem (5.8). Combining (5.10) with (5.18), we obtain the first ingredient of the proof,

$$g'_\pi(n) \ge g_\pi(n)\left(\frac{n}{\sum_{i=0}^n g_\pi(i)} - 1\right), \tag{5.36}$$

valid for $0 < \pi < 1$. The following argument is informative, if naive and incorrect. Suppose that $\pi > p_c$, and use the fact that $g_\pi(n) \to \theta(\pi)$ as $n \to \infty$, to deduce (somewhat rashly) from (5.36) that

$$\theta'(\pi) \ge 1 - \theta(\pi),$$

an inequality which is easily integrated over the interval (p_c, p) to obtain

$$\theta(p)e^p - \theta(p_c)e^{p_c} \ge e^p - e^{p_c} \qquad \text{if } p > p_c.$$

It is now a short step to deduce that

$$\theta(p) - \theta(p_c) \ge \tfrac{1}{2}\big(1 - \theta(p_c)\big)(p - p_c)$$

if $p - p_c$ is positive but sufficiently small.

A more rigorous argument proceeds as follows. The second ingredient of the proof is the fact that θ is a right-continuous function on the interval $(0, 1)$. No special technique is required to prove this; instead we use the following 'soft' argument. It is clear that $\theta(p)$ is the decreasing limit of $g_p(n)$ as $n \to \infty$. However, $g_p(n)$ is a polynomial in p and is therefore continuous, so that θ is the decreasing limit of continuous functions. Hence θ is upper semi-continuous. On the other hand, θ is non-decreasing, and is therefore right-continuous also. We shall return in Section 8.3 to the question of the continuity of θ. It turns out to be rather more difficult to show that θ is left-continuous; indeed it is an open problem to prove that θ is left-continuous at p_c, and this is equivalent to proving that $\theta(p_c) = 0$.

Let π be such that $p_c < \pi < 1$, and let ε satisfy

$$0 < \varepsilon < \tfrac{1}{2}\{1 - \theta(\pi)\}.$$

Let $\alpha \in (p_c, \pi)$. By the right-continuity of θ, we may find $\beta = \beta(\alpha) \in (\alpha, \pi]$ such that

$$\theta(\beta) \le (1+\varepsilon)\theta(\alpha). \tag{5.37}$$

Next, we pick $N = N(\alpha, \beta)$ such that

$$\frac{1}{n}\sum_{i=0}^{n} g_\beta(i) \le \frac{1}{1-\varepsilon}\theta(\beta) \qquad \text{if } n \ge N; \tag{5.38}$$

this is possible since the left side converges to $\theta(\beta)$ as $n \to \infty$. We have that

$$\theta(\alpha) \le g_\alpha(n) \le g_\gamma(n) \le g_\beta(n) \qquad \text{for all } n \text{ and } \gamma \in [\alpha, \beta],$$

giving from (5.36) that, for $n \ge N$ and $\alpha \le \gamma \le \beta$,

$$\begin{aligned} g'_\gamma(n) &\ge \theta(\alpha)\left(\frac{n}{\sum_{i=0}^{n} g_\beta(i)} - 1\right) \\ &\ge \theta(\alpha)\left(\frac{1-\varepsilon}{\theta(\beta)} - 1\right) && \text{by (5.38)} \\ &\ge \frac{1}{1+\varepsilon}\{1 - \varepsilon - \theta(\beta)\} && \text{by (5.37)} \\ &\ge \frac{1}{1+\varepsilon}\{1 - \varepsilon - \theta(\pi)\} && \text{since } \beta \le \pi \\ &\ge \frac{\varepsilon}{1+\varepsilon} && \text{since } 1 - \theta(\pi) > 2\varepsilon. \end{aligned}$$

We integrate this inequality and take the limit as $n \to \infty$ to obtain

$$\theta(\gamma) - \theta(\alpha) \ge (\gamma - \alpha)\frac{\varepsilon}{1+\varepsilon} \qquad \text{for } \alpha \le \gamma < \beta. \tag{5.39}$$

If β did not depend on the choice of α, then we would obtain the desired result by taking the limit as $\alpha \downarrow p_c$ and using the right-continuity of θ at p_c. The following argument shows that we may replace β by π in (5.39), thus completing the proof of the theorem. Define

$$\mu(\alpha) = \sup\left\{\beta \in (\alpha, \pi] : \theta(\gamma) - \theta(\alpha) \ge (\gamma - \alpha)\frac{\varepsilon}{1+\varepsilon} \text{ for all } \gamma \in [\alpha, \beta)\right\}.$$

It suffices to prove that $\mu(\alpha) = \pi$. Suppose on the contrary that $\alpha < \mu(\alpha) < \pi$. We may repeat the argument above to find that there exists $\zeta \in (\mu(\alpha), \pi]$ such that

$$\theta(\nu) - \theta(\mu(\alpha)) \geq \{\nu - \mu(\alpha)\}\frac{\varepsilon}{1+\varepsilon} \qquad \text{for } \mu(\alpha) \leq \nu < \zeta.$$

On the other hand, by the monotonicity of θ,

$$\begin{aligned}\theta(\mu(\alpha)) - \theta(\alpha) &\geq \lim_{\xi \uparrow \mu(\alpha)} \{\theta(\xi) - \theta(\alpha)\} \\ &\geq \lim_{\xi \uparrow \mu(\alpha)} \left\{(\xi - \alpha)\frac{\varepsilon}{1+\varepsilon}\right\} \\ &= \{\mu(\alpha) - \alpha\}\frac{\varepsilon}{1+\varepsilon},\end{aligned}$$

so that, for $\nu \in [\mu(\alpha), \zeta)$,

$$\begin{aligned}\theta(\nu) - \theta(\alpha) &\geq \{\nu - \mu(\alpha)\}\frac{\varepsilon}{1+\varepsilon} + \{\mu(\alpha) - \alpha\}\frac{\varepsilon}{1+\varepsilon} \\ &= (\nu - \alpha)\frac{\varepsilon}{1+\varepsilon}.\end{aligned}$$

This contradicts the maximality of $\mu(\alpha)$, and thus $\mu(\alpha) = \pi$ as claimed. □

5.3 Using Differential Inequalities

In this section we shall prove Theorem (5.2) by the method of Aizenman and Barsky (1987). Their proof is rather different from the argument in the last section, although there are certain striking similarities. The general structure of the Aizenman–Barsky proof is as follows. We introduce a new parameter $\gamma \in (0, 1)$ in such a way that ordinary percolation is obtained by taking the limit as $\gamma \downarrow 0$; this new parameter corresponds to an external magnetic field in the Ising model. For pairs (p, γ) of parameter values with $\gamma > 0$, we find that the two quantities analogous to $\theta(p)$ and $\chi(p)$ are analytic. We then establish certain differential inequalities for these quantities using Russo's formula and the BK inequality. We integrate these differential inequalities and take the limit as $\gamma \downarrow 0$ to find, as required, that θ and χ have singularities at the same value of p.

Our reasons for including this proof here are primarily twofold. First, the introduction of a new parameter is an idea with considerable physical appeal, and allows us to explore in greater depth both rigorous and intuitive links between percolation and certain models of statistical physics. The technology has other applications also. Secondly, such methods yield some information about the behaviour of θ and χ for values of p at and near the critical value p_c. This is related to the problem of 'critical exponents' to which we shall return in Chapters 9 and 10.

We introduce the new parameter γ as follows. We create a new vertex, labelled g, and we join each vertex of $\mathbb{L}^d$ by an open edge to g with probability γ, independently of all other edges. For each vertex x of $\mathbb{L}^d$, we write $\omega(x) = 1$ if x is thus joined to g and $\omega(x) = 0$ otherwise. Whereas previously the set of realizations was the set $\prod_{e\in\mathbb{E}^d}\{0,1\}$ of assignments of 'closed' or 'open' to the edges of $\mathbb{L}^d$, the set of realizations is now

$$\prod_{e\in\mathbb{E}^d}\{0,1\} \times \prod_{x\in\mathbb{Z}^d}\{0,1\},$$

since each edge is either closed or open, and in addition each vertex is either non-adjacent or adjacent to g. The original product measure $\prod_{e\in\mathbb{E}^d}\mu_e$ has been replaced by a product measure

$$\prod_{e\in\mathbb{E}^d}\mu_e \times \prod_{x\in\mathbb{Z}^d}\mu_x,$$

and we write $P_{p,\gamma}$ for this product measure and $E_{p,\gamma}$ for the corresponding expectation. We retrieve the original percolation process by neglecting g and its incident edges.

The purpose of this construction is as follows. We think of g as a surrogate 'point at infinity'. We shall concentrate on the existence of open paths joining the origin to g, and we shall take the limit as $\gamma \downarrow 0$. As γ approaches 0, the density of vertices adjacent to g drops to 0, so that the open cluster at the origin is forced to be larger and larger in order that there exist an open path from the origin to g. Writing G for the set of vertices of $\mathbb{L}^d$ which are adjacent to g, we may see that there exists an open path from the origin to g if and only if some vertex in C lies also in G; we write $\{C \cap G \neq \varnothing\}$ for this event. It is not difficult to see that, as $\gamma \to 0$,

$$P_{p,\gamma}(C \cap G \neq \varnothing) \to P_p(|C| = \infty) = \theta(p).$$

It is in this sense that ordinary percolation is retrieved by taking the limit as γ tends to 0. We shall not generally refer to the new vertex g, preferring to talk in terms of the random set G of vertices of $\mathbb{L}^d$ which are adjacent to g. The advantage of this is that the term 'open path' shall retain its meaning as a path of open edges in $\mathbb{L}^d$. Thus we shall speak in such terms as '0 is joined to some vertex in G by an open path' rather than '0 is joined to g by an open path'.

Suppose now that $0 < \gamma < 1$, and define

$$\theta(p,\gamma) = P_{p,\gamma}(C \cap G \neq \varnothing), \tag{5.40}$$

the probability that there exists an open path from the origin to some vertex in G. Just as $\theta(p)$ may be represented as

$$\theta(p) = 1 - \sum_{n=1}^{\infty} P_p(|C| = n), \tag{5.41}$$

so may we write

$$(5.42)\qquad \theta(p,\gamma) = 1 - \sum_{n=1}^{\infty} P_{p,\gamma}\big(C \cap G = \varnothing \,\big|\, |C| = n\big) P_p(|C| = n)$$
$$= 1 - \sum_{n=1}^{\infty} (1-\gamma)^n P_p(|C| = n).$$

We note that, for $\gamma > 0$, $C \cap G \neq \varnothing$ almost surely when $|C| = \infty$, and it is for this reason that the summation in (5.42) contains no term corresponding to this case. The sum in (5.42) contains non-negative terms and converges when $\gamma > 0$, and so we may take the limit as $\gamma \to 0$ to obtain that

$$(5.43)\qquad \theta(p,\gamma) \downarrow \theta(p) \qquad \text{as } \gamma \downarrow 0.$$

We define $\theta(p,0)$ to be $\theta(p)$.

The second quantity of interest is analogous to $\chi(p)$. For $0 < \gamma < 1$, we define $\chi(p,\gamma)$ to be the mean number of vertices in C on the event that there is no open path from the origin to any vertex of G; thus

$$(5.44)\qquad \chi(p,\gamma) = E_{p,\gamma}\big(|C|;\, C \cap G = \varnothing\big).$$

By conditioning on $|C|$ we see that

$$(5.45)\qquad \chi(p,\gamma) = \sum_{n=1}^{\infty} n P_{p,\gamma}\big(C \cap G = \varnothing \,\big|\, |C| = n\big) P_p(|C| = n)$$
$$= \sum_{n=1}^{\infty} n(1-\gamma)^n P_p(|C| = n).$$

As before, there is no term corresponding to the event that $|C| = \infty$, since $C \cap G \neq \varnothing$ almost surely in this case. Here also we may take the limit as $\gamma \downarrow 0$ to find that

$$(5.46)\qquad \chi(p,\gamma) \uparrow \chi^{\mathrm{f}}(p) \qquad \text{as } \gamma \downarrow 0,$$

where $\chi^{\mathrm{f}}(p) = \sum_n n P_p(|C| = n)$ is the mean size of a finite open cluster at the origin. We define $\chi(p,0)$ to be equal to $\chi^{\mathrm{f}}(p)$, and note that $\chi^{\mathrm{f}}(p) = \chi(p)$ whenever $\chi(p) < \infty$.

We see in (5.42) that θ is a power series in $1-\gamma$ having radius of convergence at least 1. We may differentiate θ with respect to γ within its disc of convergence to obtain that

$$\frac{\partial\theta}{\partial\gamma} = \sum_{n=1}^{\infty} n(1-\gamma)^{n-1} P_p(|C| = n) \qquad \text{if } 0 < \gamma < 1,$$

which we compare with (5.45) to see that

$$\chi(p,\gamma) = (1-\gamma)\frac{\partial\theta}{\partial\gamma}(p,\gamma) \qquad \text{if } 0<\gamma<1. \tag{5.47}$$

It is not difficult to show that θ is continuously differentiable with respect to p also, so long as $\gamma > 0$; to see this, it is necessary to expand $P_p(|C| = n)$ as a polynomial in p and to bound its derivative with respect to p. Rather than go through the details here, we defer this to Appendix I.

The above formulation is simple and appealing in probabilistic terms, but obscures the physical motivation underlying the parameter γ. Readers with a background in statistical physics may prefer to write $1-\gamma = e^{-h}$ where $0 \le h < \infty$, and to think of h as an external magnetic field. The quantity $\theta(p,\gamma)$ is essentially the Laplace transform of the mass function of $|C|$:

$$\theta(p,\gamma) = 1 - E\big(e^{-h|C|}\big) \qquad \text{where } h = -\log(1-\gamma) > 0.$$

In some articles, g is called the 'ghost' vertex, and vertices in G are said to be 'green'. In this jargon, $\theta(p,\gamma)$ is the probability of an open path from the origin to the ghost, and $\chi(p,\gamma)$ is the mean size of the 'green-free' open cluster at the origin.

We shall prove the following main result.

(5.48) Theorem. *If p is such that $\chi^{\mathrm{f}}(p) = \infty$ then*

either (a) $\theta(p) > 0$,

or (b) $\theta(p) = 0$ *and* $\theta(p') \ge \dfrac{1}{2p'}(p'-p)$ *for* $p' \ge p$.

This proves that $\chi(p) < \infty$ when $p < p_{\mathrm{c}}$, since suppose otherwise that $\chi(p) = \infty$ for some $p < p_{\mathrm{c}}$. For such a value of p, we have that

$$\begin{aligned} \chi^{\mathrm{f}}(p) &= E_p\big(|C|;\, |C| < \infty\big) \\ &= E_p|C| \qquad \text{since } \theta(p) = 0 \\ &= \chi(p) = \infty \end{aligned}$$

giving by the theorem that $p \ge p_{\mathrm{c}}$, a contradiction.

An important step in the proof of this theorem is the following proposition, which we extract in order that we may locate it more easily later.

(5.49) Proposition. *If p is such that $\chi^{\mathrm{f}}(p) = \infty$, there exists a constant $\alpha = \alpha(p) > 0$ such that*

$$\theta(p,\gamma) \ge \alpha\gamma^{1/2} \tag{5.50}$$

for all small positive values of γ.

Our tactics will involve using Russo's formula and the BK inequality to establish certain differential inequalities for θ and its derivatives $\partial\theta/\partial p$ and $\partial\theta/\partial\gamma = (1-\gamma)^{-1}\chi$. Unlike the differential inequality (5.23) in the proof of the last section, we shall obtain inequalities which may be integrated directly to give the result.

(5.51) Lemma. *If* $0 < \gamma < 1$ *and* $0 < p < 1$,

$$(1-p)\frac{\partial\theta}{\partial p} \le 2d(1-\gamma)\theta\frac{\partial\theta}{\partial\gamma}. \tag{5.52}$$

(5.53) Lemma. *If* $0 < \gamma < 1$ *and* $0 < p < 1$,

$$\theta \le \gamma\frac{\partial\theta}{\partial\gamma} + \theta^2 + p\theta\frac{\partial\theta}{\partial p}. \tag{5.54}$$

As remarked above, these inequalities are fairly straightforward applications of Russo's formula and the BK inequality. There is, however, an unfortunate technicality in their proofs, arising from the fact that these techniques may be applied only to events which depend on finitely many edges, whereas events such as $\{C \cap G \ne \varnothing\}$ depend on infinitely many edges. Thus it will be necessary to express such events as the limits of events depending on only finitely many edges; in physical jargon, we need to perform a finite-volume approximation. Certain properties of differentiability of θ are necessary in order to take the infinite-volume limit with rigour.

In advance of proving Lemmas (5.51) and (5.53), we show how they may be used to demonstrate Theorem (5.48) and Proposition (5.49).

Proof of Proposition (5.49). Suppose that $0 < \gamma < 1$, $0 < p < 1$, and $\chi^{\mathrm{f}}(p) = \infty$. If $\theta(p) = \theta(p, 0) > 0$ then (5.50) holds vacuously, and so we suppose that $\theta(p) = 0$. With this fixed value of p, we write $\theta(p,\gamma) = f(\gamma)$, where f is strictly increasing and continuously differentiable on $(0, 1)$ with $f(0) = 0$ and $f(1) = 1$. By the mean value theorem, for each $\gamma \in (0, 1)$, there exists $\psi = \psi(\gamma)$ satisfying $\psi \in (0,\gamma)$ such that

$$\frac{1}{\gamma} f(\gamma) = f'(\psi).$$

As $\gamma \downarrow 0$, we have that $\psi \to 0$, giving by (5.47) and (5.46) that

$$\lim_{\gamma\downarrow 0}\frac{\gamma}{f(\gamma)} = 0. \tag{5.55}$$

Let g be the inverse function of f; note that g is strictly increasing and continuously differentiable with $g(0) = 0$ and $g(1) = 1$, and satisfies

$$\lim_{\varphi\downarrow 0}\frac{1}{\varphi}g(\varphi) = 0 \tag{5.56}$$

by (5.55).

We substitute (5.52) into (5.54) to find that

$$f \le \gamma\frac{df}{d\gamma} + f^2 + 2d(1-\gamma)\frac{p}{1-p}f^2\frac{df}{d\gamma}.$$

Next we substitute $\varphi = f(\gamma)\ (> 0)$, $\gamma = g(\varphi)$, and

$$\frac{dg}{d\varphi} = \left(\frac{df}{d\gamma}\right)^{-1},$$

to obtain

$$\frac{1}{\varphi}\frac{dg}{d\varphi} - \frac{1}{\varphi^2}g \le 2d(1-g)\frac{p}{1-p} + \frac{dg}{d\varphi}.$$

By (5.56) and the preceding remarks, $g(\varphi) \ge 0$ and $g'(\varphi)$ is bounded on $(0, \Phi)$ for some $\Phi > 0$, whence there exists Φ satisfying $0 < \Phi < 1$ and a positive constant $\beta = \beta(p)$ such that

$$\frac{1}{\varphi}\frac{dg}{d\varphi} - \frac{1}{\varphi^2}g \le \beta \qquad \text{if } 0 < \varphi < \Phi.$$

We integrate this equation from $\varphi = 0$ to $\varphi = x$ where $x \le \Phi$ to obtain

$$\left[\frac{1}{\varphi}g(\varphi)\right]_0^x \le \beta x;$$

we now use (5.56) to deduce that $g(x) \le \beta x^2$ for all $x \le \Phi$. We substitute back in terms of f to obtain $\gamma \le \beta f(\gamma)^2$ for $\gamma \le g(\Phi)$, or

$$\theta(p, \gamma) = f(\gamma) \ge \alpha\gamma^{1/2}$$

where $\alpha = \beta^{-1/2}$. □

Proof of Theorem (5.48). Let $0 < \gamma < 1$, $0 < a < 1$, and suppose that $\chi^{\mathrm{f}}(a) = \infty$. If $\theta(a) > 0$, there is nothing to prove. We assume henceforth that $\theta(a) = 0$, but note that $\theta(a, \gamma) > 0$. We shall show that $\theta(b) > 0$ if $b > a$ by integrating inequality (5.54), which we write in the form

$$0 \le \frac{1}{\theta}\frac{\partial\theta}{\partial\gamma} + \frac{1}{\gamma}\frac{\partial}{\partial p}(p\theta - p). \tag{5.57}$$

Suppose $a < b < 1$. We integrate (5.57) over the rectangle $a \le p \le b, \delta \le \gamma \le \varepsilon$, where $0 < \delta < \varepsilon$; see Figure 5.4. On this rectangle, θ takes its minimum value at $(p, \gamma) = (a, \delta)$ and its maximum value at $(p, \gamma) = (b, \varepsilon)$. Thus

$$0 \le (b-a)\log\left(\frac{\theta(b,\varepsilon)}{\theta(a,\delta)}\right) + \left\{b\theta(b,\varepsilon) - b + a\right\}\log(\varepsilon/\delta),$$

where we have used the fact that $\theta(a, \delta) \ge 0$. We shall divide by $\log(\varepsilon/\delta)$ and take the limit as $\delta \downarrow 0$. From Proposition (5.49),

$$\limsup_{\delta\to 0}\left\{\frac{\log\theta(a,\delta)}{\log\delta}\right\} \le \frac{1}{2},$$

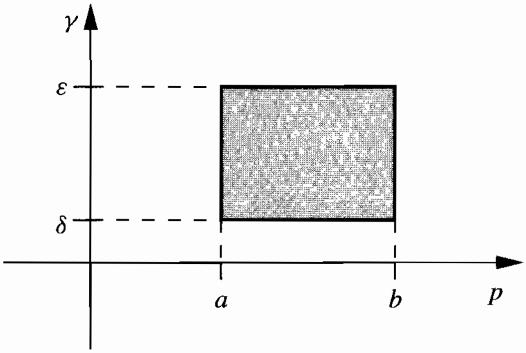


Figure 5.4. The rectangle over which we integrate the quantities in (5.57).

giving as $\delta \downarrow 0$ that

$$0 \le \tfrac{1}{2}(b-a) + b\theta(b,\varepsilon) - b + a. \tag{5.58}$$

Finally, we let ε approach 0 to obtain

$$\theta(b) = \theta(b,0) \ge \frac{1}{2b}(b-a) > 0. \qquad \square$$

We turn next to the differential inequalities (5.52) and (5.54). As remarked earlier, we shall require a finite-volume approximation. Let $B(N)$ be the cube $[-N,N]^d$ as before; we shall study the percolation process restricted to $B(N)$. It is convenient for this finite-volume process to have a property of translation invariance, and to this end we assign periodic boundary conditions to $B(N)$ by identifying opposite faces of $B(N)$. Rather than describing this in detail, we remark only that we replace by single vertices all sets of vertices of the form $\{x \in B(N) : x_i = \pm N \text{ if } i \in I, \text{ and } x_i = z_i \text{ if } i \notin I\}$ as I ranges over subsets of $\{1,2,\dots,d\}$ and z ranges over $\{-N+1,\dots,N-1\}^d$; in this new graph we identify parallel edges, leaving all other edges untouched, so that each composite vertex has degree $2d$ exactly. Thus in two dimensions we may think of $B(N)$ as being embedded in the torus. We denote the resulting graph by $L(N)$, writing $B^*(N)$ and $E(N)$ for the sets of its vertices and edges, respectively. We examine each edge of $L(N)$ in turn and declare it to be open with probability p, and closed otherwise, independently of all other edges; we then examine each vertex of $L(N)$ in turn, and join this vertex to g by an open edge with probability γ, independently of all edges and of all other vertices. Writing G_N for the set of vertices of $L(N)$ which are adjacent to g, we define quantities analogous to $\theta(p,\gamma)$ and $\chi(p,\gamma)$ by

$$\theta_N(p,\gamma) = P_{p,\gamma}(C_N \cap G_N \neq \varnothing), \tag{5.59}$$

$$\chi_N(p,\gamma) = E_{p,\gamma}\big(|C_N|;\, C_N \cap G_N = \varnothing\big), \tag{5.60}$$

where $C_N(x)$ is the open cluster of $L(N)$ at the vertex x, and $C_N = C_N(0)$. That is to say,

$$\theta_N(p,\gamma) = 1 - \sum_{n=1}^{\infty}(1-\gamma)^n P_p(|C_N| = n), \tag{5.61}$$

$$\chi_N(p,\gamma) = \sum_{n=1}^{\infty} n(1-\gamma)^n P_p(|C_N| = n). \tag{5.62}$$

These two series are finite, since $|B^*(N)| < \infty$, so that θ_N and χ_N are analytic functions of p and γ. Just as in (5.47),

$$\chi_N(p,\gamma) = (1-\gamma)\frac{\partial \theta_N}{\partial \gamma}(p,\gamma). \tag{5.63}$$

In the limit as $N \to \infty$, it is the case that $P_p(|C_N| = n) \to P_p(|C| = n)$; this rather weak remark is easy to check, since if $N > n$ and $|C_N| = n$ then there exists no vertex on the boundary of $B(N)$ which is in C_N also. Actually, rather more is true about the convergence of $P_p(|C_N| = n)$. By expanding this probability as the sum over animals of appropriate sizes, it may be shown without undue difficulty that there is sufficient regularity in such probabilities and their derivatives to ensure that $\theta_N(p,\gamma)$ and its derivatives converge as $N \to \infty$ to the appropriate derivatives of $\theta(p,\gamma)$ whenever $0 < \gamma < 1$. We do not prove this here, but summarize the argument in Appendix I. We shall therefore assume that the following limits are valid whenever $0 < \gamma < 1$ and $0 < p < 1$:

$$\theta_N(p,\gamma) \to \theta(p,\gamma) \qquad \text{as } N \to \infty, \tag{5.64}$$

$$\frac{\partial \theta_N}{\partial p}(p,\gamma) \to \frac{\partial \theta}{\partial p}(p,\gamma) \qquad \text{as } N \to \infty, \tag{5.65}$$

$$\frac{\partial \theta_N}{\partial \gamma}(p,\gamma) \to \frac{\partial \theta}{\partial \gamma}(p,\gamma) \qquad \text{as } N \to \infty. \tag{5.66}$$

Having made these assumptions, it remains only to prove (5.52) and (5.54) with θ_N in place of θ.

Proof of Lemma (5.51). Suppose $0 < \gamma < 1$ and $0 < p < 1$. We shall apply Russo's formula to the event $\{C_N \cap G_N \neq \varnothing\}$. One way of doing this is to condition on G_N, say $G_N = \Gamma$, and to apply Russo's formula to the increasing event $A_N(\Gamma) = \{C_N \cap \Gamma \neq \varnothing\}$. Thus

$$\frac{d}{dp} P_p(A_N(\Gamma)) = \sum_e P_p\big(e \text{ is pivotal for } A_N(\Gamma)\big),$$

where the sum is over all edges e of $L(N)$. The edge $e = \langle x, y\rangle$ is pivotal for $A_N(\Gamma)$ if and only if, in the graph $L(N)$ with e deleted, the following three events occur:

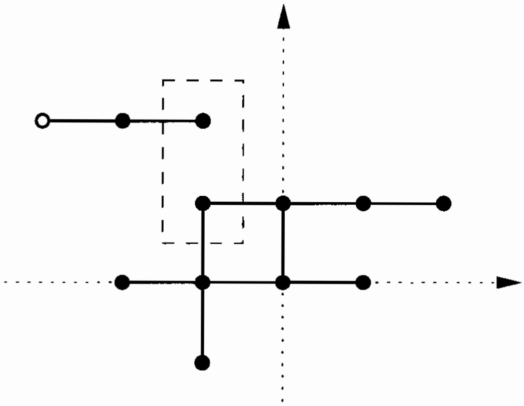


Figure 5.5. The open circle is a vertex in Γ, and the dense circles are vertices not in Γ. The origin is joined by an open path to a vertex in Γ if the two vertices in the dotted box are joined by an open edge, and not otherwise. That is to say, the edge in the dotted box is pivotal for the event $A_N(\Gamma)$.

(a) no vertex in Γ is joined to the origin by an open path;

(b) exactly one of x and y is joined to the origin by an open path; and

(c) the other endvertex of e is joined to a vertex of Γ by an open path.

See Figure 5.5 for a pictorial representation of the event that e is pivotal for $A_N(\Gamma)$. We may now see that

$$(1-p)\frac{d}{dp}P_p(A_N(\Gamma)) = \sum_e P_p(e \text{ is closed})P_p\big(e \text{ is pivotal for } A_N(\Gamma)\big)$$
$$= \sum_{x\sim y} P_p\big(x \in C_N,\ C_N \cap \Gamma = \varnothing,\ C_N(y)\cap\Gamma \neq \varnothing\big),$$

where the sum is over all ordered pairs x, y of adjacent vertices in $L(N)$. We now average over Γ, noting that

$$\sum_\Gamma (1-p)P_{p,\gamma}(G_N=\Gamma)\frac{d}{dp}P_p(A_n(\Gamma))$$
$$= (1-p)\frac{\partial}{\partial p}\left\{\sum_\Gamma P_{p,\gamma}\big(C_N\cap G_N \neq \varnothing \,\big|\, G_N=\Gamma\big)P_{p,\gamma}(G_N=\Gamma)\right\}$$
$$= (1-p)\frac{\partial \theta_N}{\partial p}$$

since G_N takes one of only finitely many values, and its distribution depends on γ only. Therefore,

$$(5.67)\quad (1-p)\frac{\partial\theta_N}{\partial p} = \sum_{x\sim y} P_{p,\gamma}\big(x\in C_N,\ C_N\cap G_N=\varnothing,\ C_N(y)\cap G_N\neq\varnothing\big).$$

The last event occurs only if $C_N(y)$ has no vertex in common with C_N. In order to bound the last probability, we condition on C_N, say $C_N = \Sigma$, and rewrite (5.67) as

$$(1-p)\frac{\partial \theta_N}{\partial p} = \sum_{x\sim y}\sum_{\Sigma} P_p(C_N = \Sigma) \times P_{p,\gamma}\Big(C_N \cap G_N = \varnothing,\ C_N(y)\cap G_N \neq \varnothing \,\Big|\, C_N = \Sigma\Big),$$

where the second sum is over all sets Σ of vertices which contain 0 and x but do not contain y. However, conditional on $C_N = \Sigma$, the events $\{C_N \cap G_N = \varnothing\}$ and $\{C_N(y) \cap G_N \neq \varnothing\}$ are independent, the first depending only on the vertices of Σ and the second depending only on the vertices outside Σ and the edges of $L(N)$ having no endvertex in Σ. Furthermore, using the translation invariance of $L(N)$,

$$\begin{aligned} P_{p,\gamma}\big(C_N(y)\cap G_N \neq \varnothing \,\big|\, C_N = \Sigma\big) &\le P_{p,\gamma}\big(C_N(y)\cap G_N \neq \varnothing\big) \\ &= \theta_N(p,\gamma), \end{aligned}$$

since the condition that $C_N = \Sigma$ where $y \notin \Sigma$ limits the possibilities for there to be an open path from y to a vertex in G_N. We deduce that

$$\begin{aligned} (1-p)\frac{\partial \theta_N}{\partial p} &\le \sum_{x\sim y}\sum_{\Sigma} P_p(C_N = \Sigma) P_{p,\gamma}\big(C_N \cap G_N = \varnothing \,\big|\, C_N = \Sigma\big)\theta_N \\ &= \theta_N \sum_{x\sim y} P_{p,\gamma}\big(x \in C_N,\ y \notin C_N,\ C_N \cap G_N = \varnothing\big) \\ &\le \theta_N \sum_{x\sim y} P_{p,\gamma}\big(x \in C_N,\ C_N \cap G_N = \varnothing\big) \\ &\le 2d\theta_N \sum_{x} P_{p,\gamma}\big(x \in C_N,\ C_N \cap G_N = \varnothing\big) \\ &\qquad\qquad \text{since } x \text{ has } 2d \text{ neighbours} \\ &= 2d\theta_N E_{p,\gamma}\big(|C_N|;\ C_N \cap G_N = \varnothing\big) \\ &= 2d\theta_N \chi_N \end{aligned}$$

as required, by (5.63). □

Proof of Lemma (5.53). This is rather similar to the last proof. Suppose that $0 < \gamma < 1$ and $0 < p < 1$, and note that

(5.68) $$\begin{aligned} \theta_N &= P_{p,\gamma}(C_N \cap G_N \neq \varnothing) \\ &= P_{p,\gamma}\big(|C_N \cap G_N| = 1\big) + P_{p,\gamma}\big(|C_N \cap G_N| \ge 2\big). \end{aligned}$$

The first term here is easily calculated by conditioning on $|C_N|$:

(5.69) $$\begin{aligned} P_{p,\gamma}\big(|C_N \cap G_N| = 1\big) &= \sum_{n=1}^{\infty} n\gamma(1-\gamma)^{n-1} P_p(|C_N| = n) \\ &= \frac{\gamma}{1-\gamma}\chi_N = \gamma\frac{\partial \theta_N}{\partial \gamma} \qquad \text{by (5.63).} \end{aligned}$$

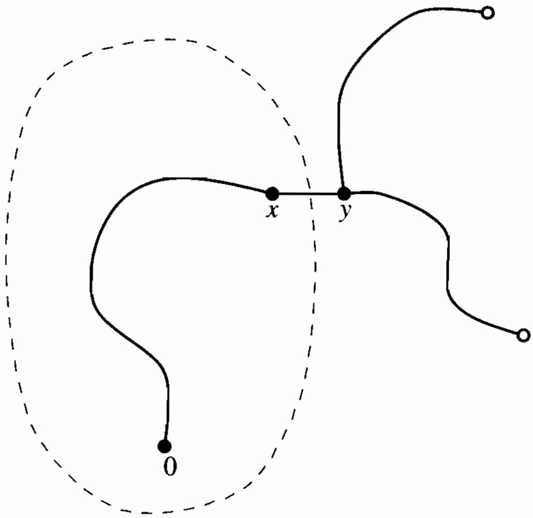


Figure 5.6. A sketch of the event that $|C_N \cap G_N| \geq 2$ but $A_0 \circ A_0$ does not occur. The open circles represent vertices in G_N, and the dashed line represents a 'hypersurface' which would cut off the origin from all vertices in G_N if $e = \langle x, y \rangle$ were closed.

The last term in (5.68) is harder to estimate. For each vertex x of $L(N)$ we write A_x for the event that: either $x \in G_N$ or x is joined by an open path to some vertex in G_N; we note that A_x is a subset of the set

$$\prod_{e \in E(N)} \{0, 1\} \times \prod_{y \in B^*(N)} \{0, 1\}$$

of realizations. Remembering the discussion leading to the BK inequality in Section 2.3, we write $A_x \circ A_x$ for the set of such realizations ω for which there exist disjoint sets V_1 and V_2 of vertices in G_N and disjoint sets E_1 and E_2 of open edges of $L(N)$, such that A_x occurs for all realizations ω' which agree with ω on (V_1, E_1) and for all realizations ω'' which agree with ω on (V_2, E_2). Thus, $A_x \circ A_x$ is the event that there exist two distinct vertices in G_N and two edge-disjoint open paths joining these two vertices to x; if $x \in G_N$, one of these paths may be the singleton x itself. Now,

(5.70)
$$P_{p,\gamma}(|C_N \cap G_N| \geq 2) = P_{p,\gamma}(A_0 \circ A_0) + P_{p,\gamma}\big(|C_N \cap G_N| \geq 2,\ A_0 \circ A_0 \text{ does not occur}\big).$$

By the BK inequality,

(5.71)
$$P_{p,\gamma}(A_0 \circ A_0) \leq P_{p,\gamma}(A_0)^2 = \theta_N^2.$$

The remaining term in (5.70) is the probability that $|C_N \cap G_N| \geq 2$ but there do not exist two disjoint connections from the origin to G_N in the sense described above. It is not hard to convince oneself (possibly with the aid of Figure 5.6) that this event occurs if and only if there exists a pair x, y of adjacent vertices of $L(N)$ such that:

(i) the edge $e = \langle x, y \rangle$ is open; and

(ii) in the subgraph of $L(N)$ obtained by deleting e, the following three events occur:

(a) no vertex of G_N is joined to the origin by an open path,

(b) x is joined to the origin by an open path,

(c) the event $A_y \circ A_y$ occurs.

The event (ii) is independent of the state of e, so that the probability that both (i) and (ii) occur is exactly

$$pq^{-1} P_{p,\gamma}\big(e \text{ is closed},\ x \in C_N,\ C_N \cap G_N = \varnothing,\ A_y \circ A_y\big),$$

where $p + q = 1$. It follows that the probability that such a pair x, y exists is no larger than

$$pq^{-1} \sum_{x \sim y} P_{p,\gamma}\big(x \in C_N,\ C_N \cap G_N = \varnothing,\ A_y \circ A_y\big),$$

since e is necessarily closed if $C_N(y)$ intersects G_N but $C_N(x)$ does not. We now condition on C_N as usual to obtain

(5.72)

$$\begin{aligned} &P_{p,\gamma}\big(|C_N \cap G_N| \geq 2,\ A_0 \circ A_0 \text{ does not occur}\big) \\ &\qquad \leq pq^{-1} \sum_{x \sim y} \sum_{\Sigma} P_p(C_N = \Sigma) P_{p,\gamma}\Big(C_N \cap G_N = \varnothing,\ A_y \circ A_y \,\Big|\, C_N = \Sigma\Big), \end{aligned}$$

where the second sum is over all sets Σ of vertices which contain 0 and x but not y. Conditional on the event $\{C_N = \Sigma\}$, the events $\{C_N \cap G_N = \varnothing\}$ and $A_y \circ A_y$ are independent, since $A_y \circ A_y$ is defined on the vertices of $L(N)$ outside Σ together with the edges between such vertices. Conditional on $\{C_N = \Sigma\}$, the states of these vertices and edges are chosen according to the appropriate product measure, implying that

(5.73)

$$
\begin{aligned}
P_{p,\gamma}\Big(C_N \cap G_N = \varnothing,\ A_y \circ A_y \,\Big|\, C_N = \Sigma\Big) \\
&= P_{p,\gamma}\big(C_N \cap G_N = \varnothing \,\big|\, C_N = \Sigma\big) P_{p,\gamma}\big(A_y \circ A_y \,\big|\, C_N = \Sigma\big) \\
&\qquad\qquad \text{by independence} \\
&\le P_{p,\gamma}\big(C_N \cap G_N = \varnothing \,\big|\, C_N = \Sigma\big) P_{p,\gamma}\big(A_y \,\big|\, C_N = \Sigma\big)^2 \\
&\qquad\qquad \text{by the BK inequality} \\
&\le P_{p,\gamma}\big(C_N \cap G_N = \varnothing \,\big|\, C_N = \Sigma\big) P_{p,\gamma}\big(A_y \,\big|\, C_N = \Sigma\big) P_{p,\gamma}(A_y) \\
&= P_{p,\gamma}\big(C_N \cap G_N = \varnothing,\, A_y \,\big|\, C_N = \Sigma\big)\theta_N \\
&\qquad\qquad \text{by independence,}
\end{aligned}
$$

where we have used the fact that A_y is more likely to occur on the whole of $L(N)$ than on the restricted graph obtained by removing Σ.

We substitute this last equation into (5.72) to obtain

(5.74)

$$
\begin{aligned}
P_{p,\gamma}\big(|C_N \cap G_N| \ge 2,\ A_0 \circ A_0 \text{ does not occur}\big) \\
&\le pq^{-1}\theta_N \sum_{x\sim y} \sum_{\Sigma} P_{p,\gamma}\big(C_N \cap G_N = \varnothing,\ C_N(y) \cap G_N \ne \varnothing,\ C_N = \Sigma\big) \\
&= pq^{-1}\theta_N \sum_{x\sim y} P_{p,\gamma}\big(x \in C_N,\ C_N \cap G_N = \varnothing,\ C_N(y) \cap G_N \ne \varnothing\big) \\
&= p\theta_N \frac{\partial \theta_N}{\partial p} \qquad \text{by (5.67),}
\end{aligned}
$$

and the result is proved by combining (5.68)–(5.71) and (5.74). □

5.4 Notes

Section 5.1. Some of the historical romance of percolation theory has been associated with the problem of proving $p_c = \frac{1}{2}$ for (bond) percolation on $\mathbb{L}^2$, and this was finally proved by Kesten (1980a), who built upon arguments of Harris (1960), Russo (1978), and Seymour and Welsh (1978). Earlier, Sykes and Essam (1964) developed a beautiful intuitive argument to show $p_c(2) = \frac{1}{2}$ (amongst other things), although their argument remains somewhat in the air even at the time of the second edition of this book (see Section 11.2). The main step in Kesten's proof was the demonstration that $p_T = p_c$ for the square lattice $\mathbb{L}^2$, where $p_T = \sup\{p : \chi(p) < \infty\}$ as in (5.3); combined with the Russo–Seymour–Welsh result that $p_T + p_c = 1$, this yielded $p_c = \frac{1}{2}$. Later, Kesten (1982) generalized his proof that $p_T = p_c$ to certain families of two-dimensional lattices; his methods were good for the plane only.

The first proof that $p_T = p_c$ in any number of dimensions was announced by Menshikov in Moscow in 1985, and appeared in Menshikov (1986) and Menshikov, Molchanov, and Sidorenko (1986). The second proof was announced in Minneapolis in early 1986 by Aizenman and Barsky (1987). It appears that Menshikov's proof first flew over the Soviet border in September 1986.

Once we know that $p_T = p_c$ in two dimensions, it is not difficult to deduce that $p_c = \frac{1}{2}$, and we shall return to this point in Chapter 11.

The following remark is of some historical relevance but little importance. Many authors have denoted the critical probability p_c by the symbol p_H; the notation p_H and p_T was introduced by Seymour and Welsh (1978) in honour of Hammersley and Temperley (see Temperley and Lieb (1971)).

Section 5.2. Menshikov (1986) reached a marginally weaker conclusion than that of Theorem (5.4), but in a slightly more general context. He was interested in finite-range percolation on a certain class of graphs satisfying conditions of regularity on their automorphism groups. When applied to the lattice $\mathbb{L}^d$, his result amounts to the statement that $P_p(A_n) < \exp\{-\psi(p)n/\log n\}$ for some $\psi(p) > 0$. The proof given here is an improved version. The renewal argument of Lemma (5.17) was found by Kesten, with a prompt from Gallavotti. Much of Menshikov's work is based upon a weaker hypothesis about the underlying graph than that it be a lattice in the conventional sense. Crucial is his assumption that the number of vertices within distance n of any given vertex is at most $\exp(n^\gamma)$ for some $\gamma < 1$. Grigorchuk (1983) has shown the existence of graphs which satisfy Menshikov's hypotheses but for which the number of such vertices grows faster than any power of n; such graphs are of substantial interest to algebraists.

The linear lower bound for $\theta(p)$ was obtained first by Chayes and Chayes (1986c, 1987a). Our argument is a variation of that of Menshikov (1987b).

Hammersley (1957a) proved a weaker form of Theorem (5.4), valid for values of p satisfying $\chi(p) < \infty$ rather then $p < p_c$; see Theorem (6.1).

Section 5.3. Ghost sites, green sites, or their equivalent, appear in Suzuki (1965), Griffiths (1967), Kasteleyn and Fortuin (1969), Fortuin and Kasteleyn (1972), and Fortuin (1972a). Differential inequalities have been used by Russo (1981, 1982), and an interesting example is in Aizenman and Newman (1984).

The Aizenman–Barsky proof is rather general; the authors show uniqueness for all finite- and long-range bond models, with or without orientations of edges, and their method applies also to finite-range site models. One of the main ideas behind Lemma (5.53) is due to J. T. Chayes and L. Chayes. Aizenman and Barsky note that their technology has close relations with technology developed for the Ising model; see Aizenman, Barsky, and Fernández (1987).

When $\gamma > 0$, the quantity analogous to the number of open clusters per vertex is

$$\kappa(p, \gamma) = \sum_{n=1}^{\infty} \frac{1}{n} (1 - \gamma)^n P_p(|C| = n).$$

We note that the function κ plays the role of the partition function of statistical mechanics in the sense that $\theta(p, \gamma)$ and $\chi(p, \gamma)$ are expressible in terms of the derivatives of κ:

$$\theta(p, \gamma) = 1 + (1 - \gamma)\frac{\partial \kappa}{\partial \gamma}, \qquad \chi(p, \gamma) = (1 - \gamma)\frac{\partial \theta}{\partial \gamma}.$$

Chapter 6

The Subcritical Phase

6.1 The Radius of an Open Cluster

In this chapter we consider the subcritical phase of bond percolation on $\mathbb{L}^d$ when $d \geq 2$; that is, we suppose that the edge-probability p satisfies $p < p_c$. In this phase, all open clusters are finite almost surely and furthermore have finite mean size. We are interested in such quantities as (i) estimates for the probability of an open path joining two vertices x and y when the distance between x and y is large, and (ii) estimates for the rate of decay of $P_p(|C| = n)$ as $n \to \infty$. Such estimates contain information about the structure of the process over long ranges, and as applications of such estimates we shall prove that $\chi(p)$ and $\kappa(p)$ are analytic functions of p when $p < p_c$.

The first estimate of this type was proved by Hammersley (1957a); he used branching process arguments to show in effect that the probability of an open path joining the origin to a vertex on the surface of the box $B(n)$ decays exponentially if $\chi(p) < \infty$. Hammersley's technique is illuminating and his result is the first and perhaps simplest in a class of such results. We begin with some notation.

As before, we write $\partial B(n)$ for the surface of $B(n)$:

$$\partial B(n) = \{x \in \mathbb{Z}^d : \|x\| = n\}$$

where $\|x\| = \max\{|x_i| : 1 \leq i \leq d\}$ as usual. For $x \in \mathbb{Z}^d$, we write $B(n, x)$ for the box $x + B(n)$ with side-length $2n$ and centre at x, and $\partial B(n, x) = x + \partial B(n)$ for the surface of this box. As usual, we write $P_p(0 \leftrightarrow \partial B(n))$ for the probability that there exists an open path joining the origin to some vertex in $\partial B(n)$.

(6.1) Theorem. Exponential tail decay of the radius of an open cluster. *Suppose that $\chi(p) < \infty$. There exists $\sigma(p) > 0$ such that*

$$P_p\big(0 \leftrightarrow \partial B(n)\big) \leq e^{-n\sigma(p)} \qquad \textit{for all } n. \tag{6.2}$$

It is largely for historical and pedagogical reasons that we state and prove Hammersley's theorem. We have already proved a stronger version of the result (Theorem (5.4)), valid under the 'milder' (but actually equivalent) condition that $p < p_c$, but the proof of Theorem (6.1) is considerably easier and is a good example of a certain type of argument.

We think of this result as demonstrating the exponential decay of the distribution of the radius of open clusters, when $\chi(p) < \infty$. We define the *radius* of the open cluster $C(x)$ at x by

$$(6.3) \qquad \text{rad}(C(x)) = \max\{\delta(x, y) : y \in C(x)\}.$$

This is the usual graph-theoretic definition of radius, although some authors (Kesten (1987e), for example) prefer to use the distance function $\|x - y\|$ in place of $\delta(x, y)$. It is easy to see that

$$(6.4) \qquad \|x - y\| \le \delta(x, y) \le d\|x - y\|,$$

so that the two definitions are equivalent within a bounded factor. We see now that

$$P_p\big(\text{rad}(C) \ge m\big) = P_p\big(0 \leftrightarrow \partial S(m)\big)$$

where $S(m) = \{y \in \mathbb{Z}^d : \delta(0, y) = m\}$ as before, giving from (6.4) and the conclusion of the theorem that

$$\begin{aligned} P_p\big(\text{rad}(C) \ge m\big) &\le P_p\big(0 \leftrightarrow \partial B(\lfloor m/d \rfloor)\big) \\ &\le \exp\big(-\sigma(p)\lfloor m/d \rfloor\big) \end{aligned}$$

when $\chi(p) < \infty$. Thus the tail of rad(C) decays at least exponentially when $\chi(p) < \infty$.

Proof of Theorem (6.1). Let $x \in \partial B(n)$, and let $\tau_p(0, x) = P_p(0 \leftrightarrow x)$ be the probability that there exists an open path of $\mathbb{L}^d$ joining the origin to x. We write N_n for the number of vertices in $\partial B(n)$ with this property, so that the mean value of N_n is

$$(6.5) \qquad E_p(N_n) = \sum_{x \in \partial B(n)} \tau_p(0, x).$$

We note that

(6.6)

$$\begin{aligned} \sum_{n=0}^{\infty} E_p(N_n) &= \sum_{n=0}^{\infty} \sum_{x \in \partial B(n)} \tau_p(0, x) \\ &= \sum_{x \in \mathbb{Z}^d} \tau_p(0, x) \qquad \text{since } (\partial B(n) : n \ge 0) \text{ partitions } \mathbb{Z}^d \\ &= E_p|\{x \in \mathbb{Z}^d : 0 \leftrightarrow x\}| = \chi(p). \end{aligned}$$

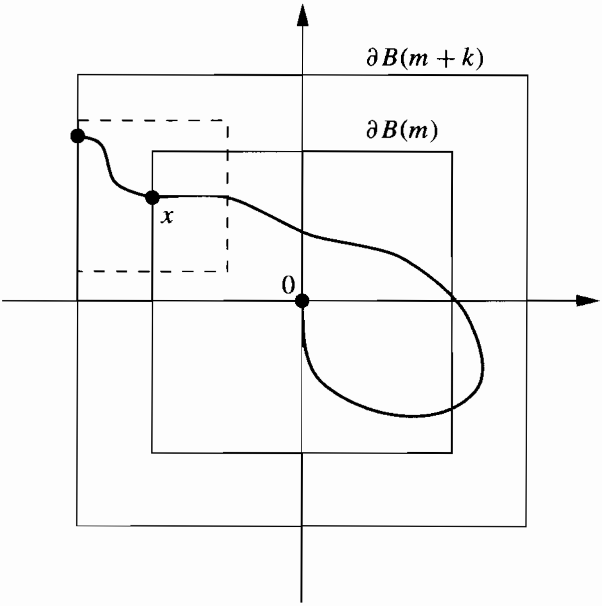


Figure 6.1. The vertex x is joined by disjoint open paths to the origin and to the surface of the box $B(k, x)$, indicated by the dashed lines.

If there exists an open path from the origin to some vertex of $\partial B(m+k)$, then there exists a vertex x in $\partial B(m)$ which is connected by disjoint open paths both to the origin and to a vertex on the surface $\partial B(k, x)$ of the box with side-length $2k$ and centre at x; see Figure 6.1. We use the BK inequality to find that

$$(6.7) \qquad P_p\big(0 \leftrightarrow \partial B(m+k)\big) \leq \sum_{x \in \partial B(m)} P_p(0 \leftrightarrow x) P_p\big(x \leftrightarrow \partial B(k, x)\big) = \sum_{x \in \partial B(m)} \tau_p(0, x) P_p\big(0 \leftrightarrow \partial B(k)\big)$$

by translation invariance. Thus

$$(6.8) \qquad P_p\big(0 \leftrightarrow \partial B(m+k)\big) \leq E_p(N_m) P_p\big(0 \leftrightarrow \partial B(k)\big) \qquad \text{for } m, k \geq 1.$$

The BK inequality makes this calculation simple; Hammersley (1957a) used a different argument.

Suppose now that $\chi(p) < \infty$, so that $\sum_{m=0}^{\infty} E_p(N_m) < \infty$ from (6.6). In this case, $E_p(N_m) \to 0$ as $m \to \infty$, and we may choose m such that $\eta = E_p(N_m)$ satisfies $\eta < 1$. Let n be a positive integer and write $n = mr + s$ where r and s

are non-negative integers and $0 \le s < m$. Then

$$\begin{aligned} P_p\big(0 \leftrightarrow \partial B(n)\big) &\le P_p\big(0 \leftrightarrow \partial B(mr)\big) && \text{since } n \ge mr \\ &\le \eta^r && \text{by iteration of (6.8)} \\ &\le \eta^{-1+n/m} && \text{since } n < m(r+1), \end{aligned}$$

which provides an exponentially decaying bound of the form of (6.2). □

The conclusion of Theorem (6.1) may be sharpened to show that, to the 'logarithmic' order, $P_p(0 \leftrightarrow \partial B(n))$ decays exactly as an exponential function of n, whenever $0 < p < p_c$. That is to say,

(6.9) $$P_p\big(0 \leftrightarrow \partial B(n)\big) \approx e^{-n\varphi(p)} \qquad \text{as } n \to \infty$$

for some $\varphi(p)$ satisfying $\varphi(p) > 0$ when $p < p_c$ (we recall that $a_n \approx b_n$ means $\log a_n / \log b_n \to 1$). Such asymptotic results are common, and the usual strategy is to use the limit theorem for subadditive sequences (see Appendix II) in order to establish the limit appropriate for (6.9), and then to use a bound such as (5.5) or (6.2) to show that the quantity $\varphi(p)$ in the exponent is strictly positive.

(6.10) Theorem. Asymptotic tail behaviour of the radius of an open cluster. *Suppose that* $0 < p \le 1$. *There exist strictly positive constants* ρ *and* σ, *independent of* p, *and a function* $\varphi(p)$, *such that*

(6.11) $$\rho n^{1-d} e^{-n\varphi(p)} \le P_p\big(0 \leftrightarrow \partial B(n)\big) \le \sigma n^{d-1} e^{-n\varphi(p)}$$

for all $n \ge 1$.

It is believed that $P_p(0 \leftrightarrow B(n))$ behaves in the manner of an exponential term multiplied by a power of n, and we return to this question in the notes at the end of this chapter.

Estimates such as (6.11) are of no value when $p > p_c$, since in this case

(6.12) $$P_p\big(0 \leftrightarrow \partial B(n)\big) \ge \theta(p) > 0 \qquad \text{for all } n;$$

thus $\varphi(p) = 0$ whenever $p > p_c$. On the other hand, $P_p(0 \leftrightarrow \partial B(n))$ decays exponentially whenever $0 < p < p_c$ (see Theorem (6.1) or Theorem (5.4)), so that $\varphi(p) > 0$ when $0 < p < p_c$. Thus

(6.13) $$\varphi(p) \begin{cases} > 0 & \text{if } 0 < p < p_c, \\ = 0 & \text{if } p > p_c. \end{cases}$$

It follows from (6.9) and (6.11) that $\varphi(p)$ is a non-increasing function of p. Other properties of φ are not so easy to ascertain, but we present some of them in the next theorem.

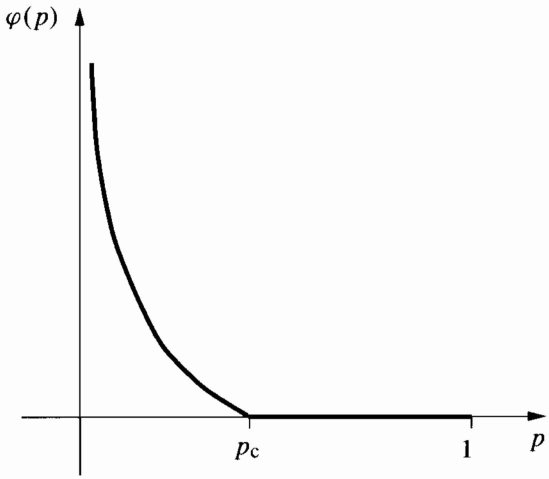


Figure 6.2. A sketch of the function φ given in Theorem (6.10).

(6.14) Theorem. *The function* φ *in* (6.11) *satisfies*:

$$\varphi \text{ is continuous and non-increasing on } (0, 1], \tag{6.15}$$

$$\varphi \text{ is strictly decreasing and positive on } (0, p_c), \tag{6.16}$$

$$\varphi(p) \to \infty \qquad \text{as } p \downarrow 0, \tag{6.17}$$

$$\varphi(p_c) = 0. \tag{6.18}$$

See Figure 6.2 for a sketch of the function φ.

We turn now to the proofs of Theorems (6.10) and (6.14). Our principal techniques are the FKG and BK inequalities, combined with the limit theorems for subadditive sequences; see Appendix II for an account of the latter technique.

Proof of Theorem (6.10). Let $0 < p \le 1$, and let $\beta(n) = P_p(0 \leftrightarrow \partial B(n))$. Our strategy is to show that the sequence $(\log \beta(n) : n \ge 1)$ is both subadditive and superadditive, subject to small error terms. The limit theorem for the subadditive inequality then delivers the goods.

Fix $m, n \ge 1$. It is not difficult to find a good upper bound for $\beta(m+n)$ involving the product $\beta(m)\beta(n)$. We recall from (6.7) and (6.8) that

$$\beta(m+n) \le \sum_{x \in \partial B(m)} \tau_p(0, x)\beta(n), \tag{6.19}$$

where $\tau_p(0, x)$ is the probability of an open path from the origin to the vertex x. However,

$$\tau_p(0, x) \le \beta(m) \qquad \text{if } x \in \partial B(m) \tag{6.20}$$

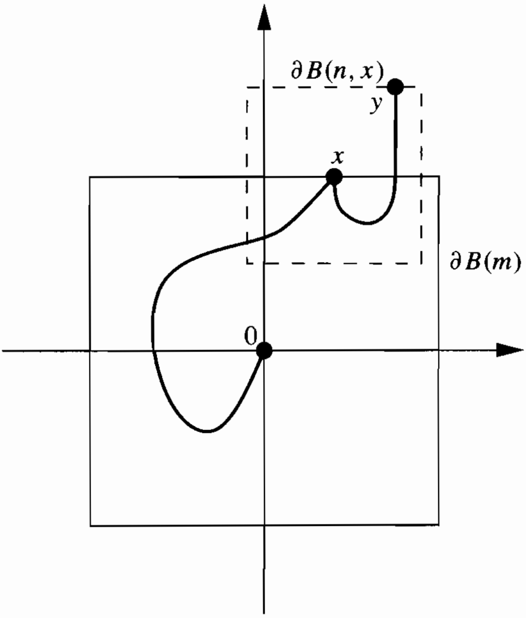


Figure 6.3. A sketch of the events (i) that x is joined by an open path of $B(m)$ to the origin, and (ii) that x is joined by an open path of $B(n, x)$ to some vertex y on the 'upper' face of $\partial B(n, x)$.

so that

$$\beta(m+n) \leq |\partial B(m)|\beta(m)\beta(n). \tag{6.21}$$

Next we establish a similar *lower* bound for $\beta(m+n)$. This is slightly more complicated and requires an estimate of the probability that there exists an open path of $B(n)$ joining the origin to a specified face on the surface of $B(n)$. More specifically, let $\gamma(n)$ be the probability that there exists some vertex $x \in \partial B(n)$ with $x_1 = n$ which is joined to the origin by an open path of $B(n)$. We have defined $\gamma(n)$ in terms of a particular face of $B(n)$, but clearly this probability is the same for each of the $2d$ faces of $B(n)$. Thus

$$\gamma(n) \leq \beta(n) \leq 2d\gamma(n). \tag{6.22}$$

Let $x \in \partial B(m)$, and choose k such that $x_k = \pm m$; we shall suppose that $x_k = m$, and an analogous argument is valid if $x_k = -m$. Let U_x be the event that there exists an open path in $B(m)$ joining x to the origin, and let V_x be the event that there exists an open path in $B(n, x)$ joining x to some vertex y of $\partial B(n, x)$ for which $y_k = m + n$; see Figure 6.3. Now,

$$\beta(m+n) \geq P_p(U_x \cap V_x) \tag{6.23}$$

since the union of two such paths contains a connection from the origin to the surface $\partial B(m+n)$. We use the FKG inequality to find that

$$\beta(m+n) \geq P_p(U_x)P_p(V_x). \tag{6.24}$$

Now, $P_p(V_x) = \gamma(n)$, and

$$\beta(m) = P_p\left(\bigcup_{x \in \partial B(m)} U_x\right) \leq \sum_{x \in \partial B(m)} P_p(U_x), \tag{6.25}$$

giving that there exists $x \in \partial B(m)$ such that

$$P_p(U_x) \geq \frac{1}{|\partial B(m)|}\beta(m). \tag{6.26}$$

With this choice of x, (6.24) becomes

$$\begin{aligned}\beta(m+n) &\geq \frac{1}{|\partial B(m)|}\beta(m)\gamma(n) \\ &\geq \frac{1}{2d|\partial B(m)|}\beta(m)\beta(n) \qquad \text{by (6.22)},\end{aligned} \tag{6.27}$$

as required.

The remaining part of the argument is devoted to analysing the consequences of inequalities (6.21) and (6.27). First, we note that

$$\begin{aligned}|\partial B(m)| &\leq 2d\left|\left\{x \in \mathbb{Z}^d : x_1 = m,\ |x_i| \leq m \text{ for } 2 \leq i \leq d\right\}\right| \\ &= 2d(2m+1)^{d-1} \\ &\leq d3^d m^{d-1} \qquad \text{for } m \geq 1,\end{aligned} \tag{6.28}$$

and then we rewrite the previous inequalities in the form

$$\log\beta(m+n) \leq \log\beta(m) + \log\beta(n) + g(m), \tag{6.29}$$

$$\log\beta(m+n) \geq \log\beta(m) + \log\beta(n) - g(m), \tag{6.30}$$

where

$$\begin{aligned}g(r) &= \log(d^2 3^{d+1} r^{d-1}) \\ &= \log(d^2 3^{d+1}) + (d-1)\log r.\end{aligned} \tag{6.31}$$

There are at least two ways of applying the subadditive limit theorem to these inequalities, and one such possibility is to note from the monotonicity of g that the sequences $(\log\beta(n) : n \geq 1)$ and $(-\log\beta(n) : n \geq 1)$ satisfy the generalized subadditive inequality with error function g. An alternative argument obviates the need to appeal to this general version of the subadditive limit theorem. Suppose that $m \le n$, and add $g(n)$ to both sides of (6.29) to obtain

$$g(n) + \log \beta(m+n) \le \{g(m) + \log \beta(m)\} + \{g(n) + \log \beta(n)\}.$$

Now, by (6.31),

$$\begin{aligned} g(m+n) - g(n) &= (d-1)\log\left(1+\frac{m}{n}\right) \\ &\le (d-1)\log 2 \end{aligned} \tag{6.32}$$

since $m \le n$. Adding these two inequalities, we find that the sequence $(a_k : k \ge 1)$, defined by

$$a_k = g(k) + (d-1)\log 2 + \log \beta(k),$$

satisfies the subadditive inequality

$$a_{m+n} \le a_m + a_n \qquad \text{for } m, n \ge 1.$$

It follows similarly from (6.30) that the sequence $\big(g(k) + (d-1)\log 2 - \log \beta(k) : k \ge 1\big)$ satisfies the subadditive inequality also. Now $k^{-1}g(k) \to 0$ as $k \to \infty$, and it follows by the subadditive limit theorem (Theorem (II.2) of Appendix II) that

$$\varphi(p) = \lim_{k\to\infty} \left\{-\frac{1}{k} \log \beta(k)\right\} \tag{6.33}$$

exists, and furthermore that

$$g(m) + (d-1)\log 2 + \log \beta(m) \ge -m\varphi(p) \tag{6.34}$$

and

$$g(m) + (d-1)\log 2 - \log \beta(m) \ge m\varphi(p) \tag{6.35}$$

for all $m \ge 1$. We combine these two inequalities to obtain

$$\begin{aligned} |m\varphi(p) + \log \beta(m)| &\le g(m) + (d-1)\log 2 \\ &\le c + (d-1)\log m \end{aligned} \tag{6.36}$$

by (6.31), for some constant $c = c(d)$. This proves (6.11). □

Proof of Theorem (6.14). Once we have taken into account the remarks preceding the statement of the theorem, it remains to prove that

(6.37) φ is continuous on $(0, 1]$, and

(6.38) φ is strictly decreasing on $(0, p_c)$, in addition to (6.17) and (6.18).

First we prove that φ is continuous on $(0, 1]$. Let

$$b_m(p) = -\frac{1}{m}\log\beta(m),$$

and write (6.36) in the form

$$|\varphi(p) - b_m(p)| \le \frac{1}{m}\big\{c + (d-1)\log m\big\}. \tag{6.39}$$

Now $b_m(p) \to \varphi(p)$ as $m \to \infty$, and this convergence is uniform on $(0, 1]$ since the right side of (6.39) does not involve p. Also, $b_m(p)$ is a continuous function of p, since the event $\{0 \leftrightarrow \partial B(m)\}$ depends only on the finite collection of edges within $B(m)$. Thus $\varphi(p)$ is the uniform limit of continuous functions, and therefore φ is continuous also. We have proved also that $\varphi(p_c) = 0$, since

$$\begin{aligned}\varphi(p_c) &= \lim_{p\downarrow p_c}\varphi(p) &&\text{by continuity}\\ &= 0 &&\text{by (6.13).}\end{aligned}$$

The strict monotonicity of φ on $(0, p_c]$ is proved by applying Theorem (2.38) to the event $\{0 \leftrightarrow \partial B(n)\}$. This is an increasing event which depends only on the edges in $B(n)$, and it is a consequence of Theorem (2.38) that

$$\frac{\log P_a(0 \leftrightarrow \partial B(n))}{\log a} \ge \frac{\log P_b(0 \leftrightarrow \partial B(n))}{\log b} \qquad \text{if } a \le b.$$

We divide by n and take the limit as $n \to \infty$ to deduce from (6.11) that

$$\frac{\varphi(a)}{\log a} \le \frac{\varphi(b)}{\log b} \qquad \text{if } 0 < a \le b \le 1;$$

hence

$$\varphi(a) \ge \varphi(b)\frac{\log(1/a)}{\log(1/b)} \qquad \text{if } 0 < a \le b \le 1. \tag{6.40}$$

Finally, we show that $\varphi(p) \to \infty$ as $p \downarrow 0$, and we use the path-counting argument of the proof of Theorem (1.10). As in (1.15) and the remarks thereafter,

$$\begin{aligned}P_p\big(0 \leftrightarrow \partial B(n)\big) &\le P_p\big(N(n) \ge 1\big)\\ &\le \big\{p\lambda(d) + \mathrm{o}(1)\big\}^n,\end{aligned} \tag{6.41}$$

where $N(n)$ is the number of open paths starting at the origin and $\lambda(d)$ is the connective constant of $\mathbb{L}^d$. We combine this with (6.11) to find that

$$\varphi(p) \ge -\log\{p\lambda(d)\} \to \infty \qquad \text{as } p \downarrow 0. \qquad \square$$

6.2 Connectivity Functions and Correlation Length

Let $x_1, x_2, \ldots, x_n$ be vertices of $\mathbb{L}^d$, and define $\tau_p(x_1, x_2, \ldots, x_n)$ to be the probability that $x_1, x_2, \ldots, x_n$ belong to the same open cluster of the percolation process. Such functions τ_p are called *connectivity functions*. We have from translation invariance that

(6.42) $$\tau_p(x_1, x_2, \ldots, x_n) = \tau_p(0, x_2 - x_1, \ldots, x_n - x_1).$$

Of particular interest is the 'two-point' connectivity function given by

(6.43) $$\tau_p(x, y) = P_p(x \leftrightarrow y).$$

In the light of Theorems (6.1) and (6.10), it is not surprising that $\tau_p(x, y)$ behaves more or less in the manner of an exponential function of the distance $\delta(x, y)$ separating x and y, when $p < p_c$. Of especial interest is the case when x is the origin and $y = e_n$ where e_n is defined by $e_n = (n, 0, \ldots, 0)$.

(6.44) Theorem. Asymptotic behaviour of the connectivity function. *Suppose that* $0 < p \leq 1$, *and let* $\varphi(p)$ *be as given in Theorem* (6.10). *Then*

(6.45) $$\lim_{n\to\infty} \left\{ -\frac{1}{n} \log \tau_p(0, e_n) \right\} = \varphi(p),$$

and there exists a strictly positive constant ζ, *independent of* p, *such that*

(6.46) $$\zeta p n^{4(1-d)} e^{-n\varphi(p)} \leq \tau_p(0, e_n) \leq e^{-n\varphi(p)} \qquad \textit{for all } n.$$

Thus $\tau_p(0, e_n)$ behaves roughly as $e^{-n\varphi(p)}$ when n is large. Some care is needed in estimating $\tau_p(0, x)$ for vertices x of $\mathbb{L}^d$ which are distant from the origin and not lying on an axis of $\mathbb{L}^d$. It is easy to adapt the argument leading to (6.45) to conclude that the limit

$$\varphi(p, x) = \lim_{n\to\infty} \left\{ -\frac{1}{n} \log \tau_p(0, nx) \right\}$$

exists for all $x \in \mathbb{Z}^d$; however, not much is known about the manner in which $\varphi(p, x)$ depends on the choice of x. On the other hand, it is not too difficult to find upper and lower bounds for $\tau_p(0, x)$ which are valid for all x.

(6.47) Proposition. *Suppose* $0 < p \leq 1$, *and let* $\varphi(p)$ *be as given in Theorem* (6.10). *There exists a strictly positive constant* λ, *independent of* p, *such that*

(6.48) $$\lambda p^d |x|^{4d(1-d)} e^{-|x|\varphi(p)} \leq \tau_p(0, x) \leq e^{-\|x\|\varphi(p)}$$

for all vertices x.

We recall that $|x| = \delta(0, x)$, $\|x\| = \max\{|x_i| : 1 \leq i \leq d\}$, and $\|x\| \leq |x| \leq d\|x\|$ for all x. Thus the inequalities in (6.48) specify the decay rate of $\tau_p(0, x)$ up to a numerical factor in the exponent which depends on the number of dimensions.

The above remarks make extensive reference to the function φ, but how may we relate φ to other relevant quantities? The following concrete upper bound on $\tau_p(0, x)$ is of some value here.

(6.49) Proposition. *Suppose* $0 < p < p_c$, *and let* $\chi(p)$ *be the mean cluster size and* $\varphi(p)$ *be as given in Theorem* (6.10). *Then*

(6.50) $$\tau_p(0, x) \le \left\{1 - \chi(p)^{-1}\right\}^{|x|} \qquad \textit{for all } x,$$

giving that

(6.51) $$\varphi(p)^{-1} \le \chi(p) \qquad \textit{if } 0 < p < p_c.$$

We recall from Theorem (6.14) that $\varphi(p) \to 0$ as $p \uparrow p_c$, and it follows from (6.51) that

(6.52) $$\chi(p) \to \infty \qquad \text{as } p \uparrow p_c.$$

We have seen that probabilities such as $P_p(0 \leftrightarrow \partial B(n))$ and $\tau_p(0, e_n)$ decay roughly as $e^{-n\varphi(p)}$ where $\varphi(p) > 0$, if $0 < p < p_c$. It is believed that such probabilities behave in the manner of a power of n multiplied by the exponential factor $e^{-n\varphi(p)}$; remember (6.11) and (6.46). [This fact is known to hold for $\tau_p(0, e_n)$, and goes by the name 'Ornstein–Zernike decay'; see the notes for this section.] On the other hand, there is no such exponential decay when $p = p_c$, since $\varphi(p_c) = 0$; actually, it is likely that such probabilities behave like negative powers of n when $p = p_c$. Thus a subcritical process with edge-probability p behaves qualitatively differently from the critical process with edge-probability p_c only when the spatial scale n on which we observe the process is such that $n\varphi(p)$ is large. This observation is particularly important when the difference between p and p_c is small, since then $\varphi(p)$ is small also, so that the two processes differ qualitatively only over a large spatial scale. As p approaches p_c from beneath, we have that $\varphi(p) \to 0$, so that the required scale becomes larger and larger. This notion of scaling seems to be central to understanding the behaviour of the percolation process at and near the critical point p_c. In such discussions it is usual to emphasize not the function $\varphi(p)$ but its reciprocal $\xi(p) = \varphi(p)^{-1}$, where we define $\xi(p_c) = \infty$ since $\varphi(p_c) = 0$; the quantity $\xi(p)$ is called the *correlation length* of the percolation process with edge-probability p, since only when n is of order $\xi(p)$ or greater do the exponential terms in expressions such as

(6.53) $$P_p\big(0 \leftrightarrow \partial B(n)\big) \approx e^{-n/\xi(p)}$$

and

(6.54) $$\tau_p(0, e_n) \approx e^{-n/\xi(p)}$$

become significant. Henceforth, we shall describe the asymptotics of such probabilities in terms of $\xi(p)$ rather than $\varphi(p)$. We note from Theorem (6.14) and equation (6.51) that ξ is continuous and strictly increasing on $(0, p_c)$, and satisfies

(6.55) $$\xi(p) \to 0 \qquad \text{as } p \downarrow 0,$$

(6.56) $$\xi(p) \to \infty \qquad \text{as } p \uparrow p_c,$$

(6.57) $$\xi(p) \le \chi(p) \qquad \text{if } 0 < p < p_c.$$

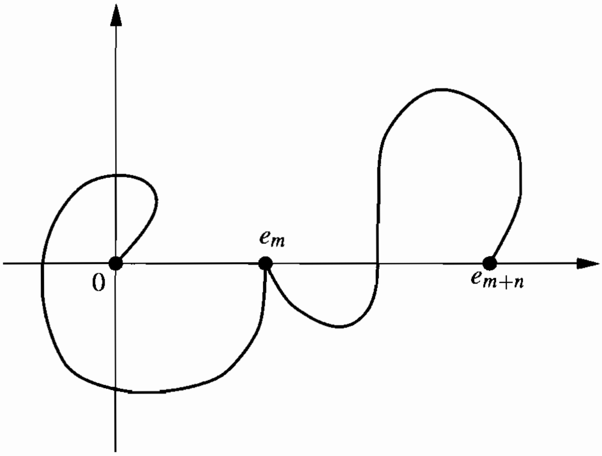


Figure 6.4. If $0 \leftrightarrow e_m$ and $e_m \leftrightarrow e_{m+n}$, then $0 \leftrightarrow e_{m+n}$.

Note also that the correlation length $\xi(p)$ is defined only for $p \le p_c$. In Chapter 8, we shall encounter the natural counterpart of ξ when $p > p_c$.

Proof of Theorem (6.44). This is similar to but simpler than the proof of Theorem (6.10). Clearly,

$$\{0 \leftrightarrow e_{m+n}\} \supseteq \{0 \leftrightarrow e_m\} \cap \{e_m \leftrightarrow e_{m+n}\}; \tag{6.58}$$

see Figure 6.4 for an illustration of this. By the FKG inequality and translation invariance,

$$\tau_p(0, e_{m+n}) \ge \tau_p(0, e_m)\tau_p(e_m, e_{m+n}) = \tau_p(0, e_m)\tau_p(0, e_n). \tag{6.59}$$

Thus $t(k) = -\log \tau_p(0, e_k)$ satisfies

$$t(m+n) \le t(m) + t(n),$$

and the subadditive limit theorem (Theorem (II.2) of Appendix II) gives us that the limit

$$\eta(p) = \lim_{n\to\infty} \left\{ -\frac{1}{n} \log \tau_p(0, e_n) \right\}$$

exists. Furthermore,

$$\log \tau_p(0, e_n) = t(n) \le -n\eta(p) \quad \text{for all } n.$$

We show next that $\eta(p) = \varphi(p)$. First note that

$$\tau_p(0, e_n) \le P_p\big(0 \leftrightarrow \partial B(n)\big),$$

as in (6.20); this implies that $\eta(p) \geq \varphi(p)$. It requires slightly more work to show the opposite inequality.

We have from (6.26) that there exists a vertex x in $\partial B(m)$ such that

$$P_p\big(0 \leftrightarrow x \text{ in } B(m)\big) \geq \frac{1}{|\partial B(m)|} P_p\big(0 \leftrightarrow \partial B(m)\big). \tag{6.60}$$

Using the rotation invariance of the lattice, we may pick such a vertex x satisfying $x_1 = m$. By symmetry, the events $\{0 \leftrightarrow x\}$ and $\{x \leftrightarrow e_{2m}\}$ are equally likely (just reflect the set of paths from 0 to x in the hyperplane $\{y \in \mathbb{Z}^d : y_1 = m\}$ to obtain the isomorphic set of paths from e_{2m} to x). On the other hand,

$$\{0 \leftrightarrow e_{2m}\} \supseteq \{0 \leftrightarrow x\} \cap \{x \leftrightarrow e_{2m}\}$$

as in (6.58). We use the FKG inequality here to obtain

$$\tau_p(0, e_{2m}) \geq \tau_p(0, x)^2, \tag{6.61}$$

which we combine with (6.60) to obtain

$$\begin{aligned}\tau_p(0, e_{2m}) &\geq A_1 m^{2(1-d)} P_p\big(0 \leftrightarrow \partial B(m)\big)^2 \\ &\geq A_2 m^{4(1-d)} e^{-2m\varphi(p)} \qquad \text{by (6.11)}\end{aligned}$$

for appropriate positive constants A_1 and A_2; we have used the fact that $|\partial B(m)|$ is of order m^{d-1}. Not only does it follow that $\eta(p) = \varphi(p)$, but also we have obtained the left inequality in (6.46) for even values of n. It is only a minor nuisance that we need this inequality for odd n also. If $n = 2m + 1$ then, with the same choice of x,

$$\{0 \leftrightarrow e_{2m+1}\} \supseteq \{0 \leftrightarrow x\} \cap \big\{\langle x, x + e_1\rangle \text{ is open}\big\} \cap \{x + e_1 \leftrightarrow e_{2m+1}\},$$

an intersection of increasing events. Applying the FKG inequality, we obtain that

$$\tau_p(0, e_{2m+1}) \geq p\tau_p(0, x)^2$$

and the result follows as before. □

Proof of Proposition (6.47). Suppose that $\|x\| = n$, so that $x \in \partial B(n)$. We have from (6.61) that

$$\tau_p(0, x) \leq \tau_p(0, e_{2n})^{1/2} \leq e^{-n\varphi(p)} \qquad \text{by (6.46)},$$

and this is the second inequality which we are required to prove.

Turning to the first inequality in (6.48), let us suppose that $|x| = n$, so that

$$\sum_{i=1}^{d} |x_i| = n.$$

Let $y_i = (|x_1|, |x_2|, \dots, |x_i|, 0, \dots, 0)$ for $0 \le i \le d$. By symmetry, $\tau_p(0, x) = \tau_p(0, y_d)$. Now,

$$\{0 \leftrightarrow y_d\} \supseteq \bigcap_{0<i\le d} \{y_{i-1} \leftrightarrow y_i\} \tag{6.62}$$

by the usual argument. By the FKG inequality and the symmetry of the process,

$$\tau_p(0, y_d) \ge \prod_{i=1}^{d} \tau_p(y_{i-1}, y_i) = \prod_{i=1}^{d} \tau_p(0, e_{|x_i|}). \tag{6.63}$$

For large $|x_i|$, $\tau_p(0, e_{|x_i|})$ behaves approximately as $\exp(-\varphi(p)|x_i|)$, whence this last product behaves approximately as

$$\exp\left(-\varphi(p) \sum_i |x_i|\right) = e^{-\varphi(p)n} \tag{6.64}$$

as required. Arguing more rigorously, it follows from (6.46) that

$$\tau_p(0, y_d) \ge Ap^d \left\{|x_1||x_2| \dots |x_d|\right\}^{4(1-d)} e^{-n\varphi(p)} \tag{6.65}$$

for some positive constant A, and the required conclusion is obtained by way of the standard inequality for geometric and arithmetic means,

$$\left\{|x_1||x_2| \dots |x_d|\right\}^{1/d} \le \frac{1}{d} \sum_{i=1}^{d} |x_i| = n/d. \qquad \square$$

Proof of Proposition (6.49). We follow Aizenman and Newman (1984). The argument is closely related to the proof of Theorem (6.1), except that we seek an inequality involving the norm $\delta(0, x) = |x|$ rather than $\|x\|$. Instead of the box $B(n)$, we consider the sphere $S(n) = \{x \in \mathbb{Z}^d : |x| \le n\}$ and its boundary $\partial S(n) = \{x \in \mathbb{Z}^d : |x| = n\}$; similarly, we write $S(n, x) = x + S(n)$ and $\partial S(n, x) = x + \partial S(n)$ for the corresponding sets centred at x. Let M_n be the number of vertices in $\partial S(n)$ which are connected to the origin by open paths, so that

$$E_p(M_n) = \sum_{x \in \partial S(n)} \tau_p(0, x) \tag{6.66}$$

as in (6.5). Similarly to (6.6), we have that

$$\sum_{n=0}^{\infty} E_p(M_n) = \chi(p). \tag{6.67}$$

The argument leading to (6.7) is easily adapted to give

$$\tau_p(0,z) \le \sum_{x\in\partial S(m)} \tau_p(0,x)\tau_p(x,z) \tag{6.68}$$

whenever $|z| \ge m$. The term $\tau_p(x,z)$ is no greater than 1, so that

$$\tau_p(0,z) \le E_p(M_m) \qquad \text{when } |z| \ge m, \tag{6.69}$$

by comparison with (6.66). We can do much better than this. We fix m and suppose that $|z| = n$. We write $n = mr+s$ for non-negative integers r, s satisfying $0 \le s < m$. We now iterate (6.68) to obtain

$$\begin{aligned}\tau_p(0,z) &\le \sum_{x\in\partial S(m)} \tau_p(0,x) \sum_{y\in\partial S(m,x)} \tau_p(x,y)\tau_p(y,z) \\ &\le E_p(M_m)^r && \text{as in (6.69)} \\ &= E_p(M_m)^{\lfloor |z|/m \rfloor} && \text{for all } z \text{ and } m.\end{aligned} \tag{6.70}$$

We may improve this still further as follows. Let $u \in \mathbb{Z}^d$ and let k be a positive integer. By the usual argument, we have that

$$\tau_p(0,ku) \ge P_p\big(ju \leftrightarrow (j+1)u \text{ for } 0 \le j < k\big) \ge \tau_p(0,u)^k$$

by the FKG inequality. Thus

$$\begin{aligned}\tau_p(0,u) &\le \tau_p(0,ku)^{1/k} \\ &\le E_p(M_m)^{\lfloor |ku|/m \rfloor /k} && \text{by (6.70)} \\ &\to E_p(M_m)^{|u|/m} && \text{as } k \to \infty\end{aligned} \tag{6.71}$$

for all vertices u and all integers m.

Suppose now that $0 < p < p_c$, so that $\chi(p) < \infty$. There exists $m \ge 1$ such that

$$E_p(M_m) \le \big\{1 - \chi(p)^{-1}\big\}^m, \tag{6.72}$$

since, if $E_p(M_m) > \{1 - \chi(p)^{-1}\}^m$ for all $m \ge 1$, then (6.67) implies that

$$\chi(p) > \sum_{m=0}^{\infty} \big\{1 - \chi(p)^{-1}\big\}^m = \chi(p),$$

a contradiction. We choose $m \ge 1$ such that $E_p(M_m) \le \{1 - \chi(p)^{-1}\}^m$ and substitute this into (6.71) to obtain

$$\tau_p(0,u) \le \big\{1 - \chi(p)^{-1}\big\}^{|u|} \tag{6.73}$$

as claimed. Inequality (6.51) is a consequence of the fact that $1 - r \le e^{-r}$ when $r \ge 0$. □

6.3 Exponential Decay of the Cluster Size Distribution

We turn our attention now to the number $|C|$ of vertices in the open cluster C at the origin. If $p < p_c$, this cluster is almost surely finite. We require an estimate on the tail of the distribution of $|C|$. It is easy to see (as in Section 5.2) that

$$P_p\big(0 \leftrightarrow \partial B(n)\big) \le P_p(|C| \ge n) \le P_p\big(0 \leftrightarrow \partial B(\alpha n^{1/d})\big) \tag{6.74}$$

for some positive constant $\alpha = \alpha(d)$; thus the exponential decay of the sequence $P_p(0 \leftrightarrow \partial B(n))$ implies that $P_p(|C| \ge n)$ decays at least as fast as $\exp(-\gamma n^{1/d})$ for some positive $\gamma = \gamma(p, d)$. This is not the best possible result, as the next theorem indicates.

(6.75) Theorem. Exponential decay of the cluster size distribution. *Suppose that* $0 < p < p_c$. *There exists* $\lambda(p) > 0$ *such that*

$$P_p(|C| \ge n) \le e^{-n\lambda(p)} \qquad \textit{for all } n \ge 1. \tag{6.76}$$

We shall actually prove a more concrete inequality than this, in which $\lambda(p)$ is expressed in terms of the mean cluster size χ. More precisely, we shall prove that

$$P_p(|C| \ge n) \le 2\exp\big(-\tfrac{1}{2}n/\chi(p)^2\big) \qquad \text{if } n > \chi(p)^2. \tag{6.77}$$

We may use subadditivity to show that $-n^{-1}\log P_p(|C| = n)$ converges as $n \to \infty$, and the inequalities above show that the limit is finite and strictly positive whenever $0 < p < p_c$.

(6.78) Theorem. *Suppose* $0 < p < 1$. *The limit*

$$\zeta(p) = \lim_{n\to\infty}\left\{-\frac{1}{n}\log P_p(|C| = n)\right\} \tag{6.79}$$

exists and satisfies $0 \le \zeta(p) < \infty$ *for all* $p \in (0, 1)$. *Also*

$$P_p(|C| = n) \le \frac{(1-p)^2}{p} n e^{-n\zeta(p)} \qquad \textit{for all } n \ge 1. \tag{6.80}$$

The limit $\zeta(p)$ *is strictly positive if* $0 < p < p_c$.

We note that

$$P_p(n \le |C| < \infty) = \sum_{m=n}^{\infty} P_p(|C| = m), \tag{6.81}$$

so that the limit ζ in (6.79) satisfies

$$\zeta(p) = \lim_{n\to\infty}\left\{-\frac{1}{n}\log P_p(n \le |C| < \infty)\right\} \tag{6.82}$$

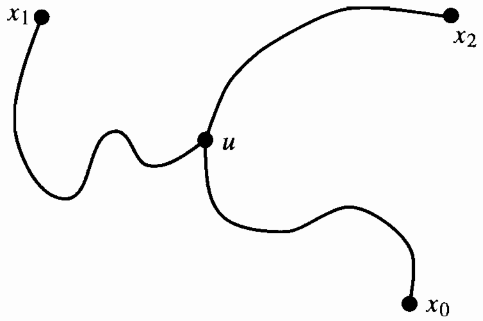


Figure 6.5. If x_0, x_1, and x_2 are in the same open cluster, there exists a vertex u which is joined by edge-disjoint open paths to x_0, x_1, and x_2.

also (this is seen easiest by considering the cases $\zeta(p) = 0$ and $\zeta(p) > 0$ separately). We have from (6.74) that

(6.83) $$\zeta(p) \leq \varphi(p) \qquad \text{for } 0 < p < p_c,$$

where φ is the decay rate of $P_p(0 \leftrightarrow \partial B(n))$ given in Theorem (6.10). We shall see in Theorem (8.61) that the tail of $|C|$ decays slower than exponentially when $p > p_c$, so that $\zeta(p) = 0$ if $p > p_c$.

Proof of Theorem (6.75). We use the method of Aizenman and Newman (1984) aided by the BK inequality. We shall obtain an exponentially decaying upper bound for $P_p(|C| \geq n)$ by showing that

$$\nu(t) = E_p\big(|C| \exp(t|C|)\big)$$

is finite for small positive values of t, and then using Markov's inequality. For $t > 0$, $\nu(t)$ is the mean of an increasing random variable, and this suggests that the BK inequality may be useful in finding an upper bound for $\nu(t)$. We shall suppose henceforth that $0 < p < p_c$ so that $\chi(p) < \infty$, and we shall concentrate on finding good upper bounds for the moments $E_p(|C|^n)$. Clearly

(6.84) $$|C| = \sum_{x \in \mathbb{Z}^d} I_{\{0 \leftrightarrow x\}},$$

the sum over x of the indicator function of the event that $x \in C$. Thus

(6.85) $$E_p(|C|^n) = E_p\left(\sum_{x_1, \dots, x_n} I_{\{0 \leftrightarrow x_1\}} I_{\{0 \leftrightarrow x_2\}} \dots I_{\{0 \leftrightarrow x_n\}}\right) = \sum_{x_1, \dots, x_n} \tau_p(0, x_1, x_2, \dots, x_n),$$

where the summations are over all n-tuples $(x_1, x_2, \dots, x_n)$ of (possibly non-distinct) vertices of $\mathbb{L}^d$, and the connectivity function τ_p is given by

$$\tau_p(y_0, y_1, \dots, y_n) = P_p\big(y_0, y_1, \dots, y_n \text{ belong to the same open cluster}\big).$$

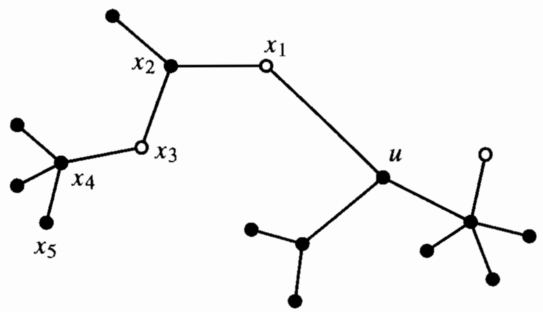


Figure 6.6. An illustration of the proof of Lemma (6.87). The open circles indicate vertices in W, so that $J = 3$ and $w = x_3$.

We need upper bounds for these connectivity functions. Note that

$$E_p(|C|^0) = 1, \qquad E_p(|C|) = \chi(p), \tag{6.86}$$

so that the first case of interest is when $n = 2$. Suppose that x_0, x_1, and x_2 are distinct vertices which are in the same open cluster of the process. We claim that there exists a vertex u (possibly equal to x_0, x_1, or x_2) such that u is joined by three edge-disjoint open paths to x_0, x_1, and x_2, respectively; if u equals one of the three original vertices, say $u = x_0$, then the required path joining u to x_0 may be taken to be the trivial path with zero length. See Figure 6.5. This observation is an elementary fact of graph theory, and is a consequence of the following lemma.

(6.87) Lemma. *Let G be a connected graph and let W be a non-empty subset of the vertex set of G. There exists a vertex $w \in W$ such that, in the graph obtained from G by deleting w and its incident edges, all vertices in $W \setminus \{w\}$ are in the same connected component.*

Proof. Let T be a spanning tree of G and let u be any given vertex of G; we think of T as being a tree with root at u. We may suppose that $|W| \geq 2$, so that at least one of the branches of T contains a vertex in W. Let $u = x_0, x_1, x_2, \ldots, x_m$ be the vertices on such a branch, listed in increasing order of distance from u in the tree T. We define $J = \max\{j : x_j \in W\}$ and set $w = x_J$. A glance at Figure 6.6 confirms that w has the required property. □

We apply this lemma with $G = C(x_0)$ and $W = \{x_0, x_1, x_2\}$, to deduce that there exists an open path π_1 joining two of the vertices of W which does not use the third such vertex. However, $C(x_0)$ is connected, so that there exists an open path π_2 from the third vertex to some vertex in π_1, using no other vertex in π_1; we may take u to be the vertex common to π_1 and π_2. We have assumed above that x_0, x_1, x_2 are distinct, but this assumption is not essential. If $x_0 = x_1 = x_2$ say, we may take $u = x_0$, and if $x_0 = x_1 \neq x_2$ we may take $u = x_0$. We have proved that

$$(6.88) \qquad \tau_p(x_0, x_1, x_2) \le \sum_{u \in \mathbb{Z}^d} P_p\big(\{u \leftrightarrow x_0\} \circ \{u \leftrightarrow x_1\} \circ \{u \leftrightarrow x_2\}\big),$$

and the BK inequality (particularly inequality (2.17)) yields the next lemma.

(6.89) Lemma. *For all values of p and vertices x_0, x_1, x_2 we have that*

$$(6.90) \qquad \tau_p(x_0, x_1, x_2) \le \sum_{u \in \mathbb{Z}^d} \tau_p(u, x_0)\tau_p(u, x_1)\tau_p(u, x_2).$$

We may apply this to (6.85) with $n = 2$ to find that

$$(6.91) \qquad \begin{aligned} E_p(|C|^2) &\le \sum_{u, x_1, x_2} \tau_p(u, 0)\tau_p(u, x_1)\tau_p(u, x_2) \\ &= \Big(\sum_y \tau_p(0, y)\Big)^3 = (E_p|C|)^3 \\ &= \chi(p)^3. \end{aligned}$$

We have treated the case $n = 2$ in some detail, since it indicates a route to an upper bound for $E_p(|C|^n)$ for general n.

We begin with some notation. A *tree* is a connected graph with no circuits; a tree on n vertices has $n - 1$ edges. We call a tree a *skeleton* if each vertex has degree either 1 or 3. It is an easy calculation to see that if a skeleton has k vertices of degree 1 (called *exterior* vertices) then it has $k - 2$ vertices of degree 3 (called *interior* vertices). We denote by $I(S)$ the set of interior vertices of the skeleton S. A skeleton with k exterior vertices is called *labelled* if there is an assignment of the numbers $0, 1, 2, \dots, k - 1$ to its exterior vertices. Two labelled skeletons S_1 and S_2 are called *isomorphic* if there is a one–one correspondence between the vertex sets of S_1 and S_2 under which both the adjacency relation and the labellings of the exterior vertices are preserved. We shall be interested in labelled skeletons, of which some examples are presented in Figure 6.7.

We saw above that, if x_0, x_1, and x_2 are in the same open cluster, there exists a labelled skeleton S with three exterior vertices together with a mapping $\psi_{\mathbf{x}}$ from the vertex set of S into $\mathbb{Z}^d$ such that the following hold:

(a) the exterior vertex of S with label i is mapped to x_i by $\psi_{\mathbf{x}}$, for $i = 0, 1, 2$; and

(b) the edges of S correspond to edge-disjoint open paths of the lattice, in that there exist three edge-disjoint open paths joining the three pairs $(\psi_{\mathbf{x}}(v), x_0)$, $(\psi_{\mathbf{x}}(v), x_1)$, and $(\psi_{\mathbf{x}}(v), x_2)$, where v is the unique interior vertex of S.

We now make this construction more general. Suppose that $x_0, x_1, \dots, x_k$ are (possibly non-distinct) vertices of $\mathbb{L}^d$ which belong to the same open cluster. We

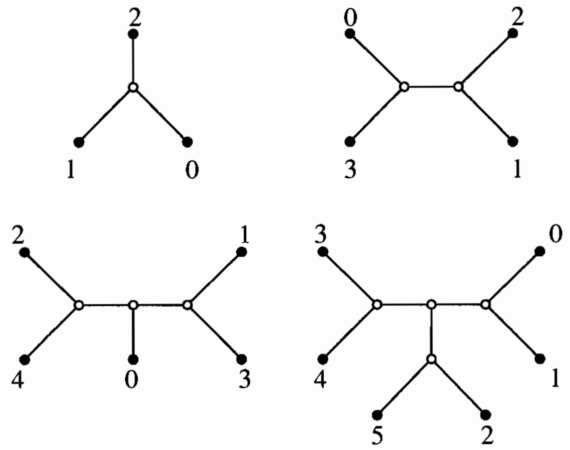


Figure 6.7. Some labelled skeletons. The open circles denote interior vertices.

claim that there exists a labelled skeleton S with $k+1$ exterior vertices together with a mapping $\psi_{\mathbf{x}}$ from the vertex set of S into $\mathbb{Z}^d$, such that the following hold:

(c) the exterior vertex of S with label i is mapped to x_i by $\psi_{\mathbf{x}}$, for $i = 0, 1, \dots, k$; and

(d) the edges of S correspond to edge-disjoint open paths of the lattice, in that there exist $2k-1$ edge-disjoint open paths joining the $2k-1$ pairs $\{(\psi_{\mathbf{x}}(u), \psi_{\mathbf{x}}(v)) : \langle u, v\rangle \in S\}$.

Here, $\langle u, v\rangle$ is the edge of S with endvertices u and v. We prove this claim by induction on k, noting that it is valid for $k = 2$. Suppose that it is valid for $k = n$ and suppose that $x_0, x_1, \dots, x_{n+1}$ belong to the same open cluster of $\mathbb{L}^d$. From Lemma (6.87), there exists j such that every vertex in $\{x_i : 0 \le i \le n+1\} \setminus \{x_j\}$ belongs to the same connected component of the graph obtained from this cluster by deleting x_j and its incident edges; we may assume without loss of generality that $j = n+1$. Suppose for the moment that none of $x_0, x_1, \dots, x_n$ equals x_{n+1}. By the induction hypothesis, there exists a labelled skeleton S with $n+1$ exterior vertices, together with a mapping $\psi_{\mathbf{x}}$, satisfying (c) and (d) above with $k = n$, such that x_{n+1} lies in none of the $2n-1$ edge-disjoint open paths referred to in (d). However, x_{n+1} is in the same open cluster as $x_0, x_1, \dots, x_n$, and so there exists an open path π joining x_{n+1} to some vertex in these $2n-1$ edge-disjoint paths, using exactly one vertex (z, say) in these paths. We amend the skeleton S accordingly (see Figure 6.8), adding an additional interior vertex v at the appropriate place, and extending the domain of the mapping $\psi_{\mathbf{x}}$ by defining $\psi_{\mathbf{x}}(v) = z$; thus the claim is valid for $k = n+1$ also. We have assumed that x_{n+1} appears only once in the sequence $x_0, x_1, \dots, x_{n+1}$, but this assumption is in no way essential. Suppose, for example, that x_{n+1} appears twice and $x_0 = x_{n+1}$, say. The above argument remains valid, and in this case π is the path containing the single vertex x_{n+1} ($= x_0 = z$) and no edges.

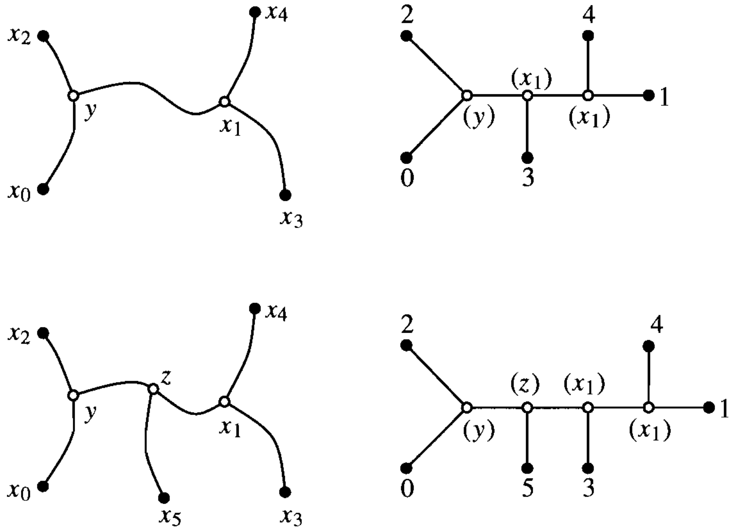


Figure 6.8. The upper pair of pictures depict interconnections between x_0, x_1, x_2, x_3, x_4 together with an associated labelled skeleton S; the labels in parentheses are the images of the appropriate vertices of S under $\psi_{\mathbf{x}}$. In the lower pair of pictures, the vertex x_5 has been adjoined.

It follows that, for $k \geq 2$,

$$\tau_p(x_0, x_1, \dots, x_k) \leq \sum_S \sum_{\psi_{\mathbf{x}}} P_p\big(\text{there exist edge-disjoint paths joining } \psi_{\mathbf{x}}(u) \text{ to } \psi_{\mathbf{x}}(v) \text{ for each edge } \langle u, v\rangle \text{ of } S\big), \tag{6.92}$$

where the first sum is over all labelled skeletons with $k+1$ exterior vertices and the second sum is over all admissible mappings $\psi_{\mathbf{x}}$ from the vertex set of S into $\mathbb{Z}^d$. The second sum amounts to the summation over all possible choices for $\psi_{\mathbf{x}}(v)$ as v ranges over $I(S)$, the set of interior vertices of S. The above inequality reduces to (6.88) when $k = 2$. We use the BK inequality to find that

$$\tau_p(x_0, x_1, \dots, x_k) \leq \sum_S \sum_{\psi_{\mathbf{x}}} \prod_{\langle u,v\rangle \in S} \tau_p\big(\psi_{\mathbf{x}}(u), \psi_{\mathbf{x}}(v)\big), \tag{6.93}$$

in agreement with (6.90). Hence, for $n \geq 2$,

$$\begin{aligned} E_p(|C|^n) &= \sum_{x_1, x_2, \dots, x_n} \tau_p(x_0, x_1, x_2, \dots, x_n) \\ &\leq \sum_{x_1, x_2, \dots, x_n} \sum_S \sum_{\psi_{\mathbf{x}}} \prod_{\langle u,v\rangle \in S} \tau_p\big(\psi_{\mathbf{x}}(u), \psi_{\mathbf{x}}(v)\big), \end{aligned}$$

by (6.85) and (6.93); we have written $x_0 = 0$. When we perform the summation

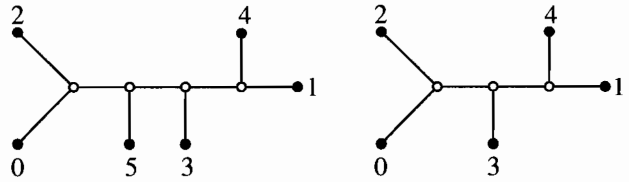


Figure 6.9. A labelled skeleton with six exterior vertices, and the skeleton obtained by deleting the vertex labelled 5 together with the incident edge. The smaller skeleton has five exterior vertices and seven edges, and it is not difficult to check that the seven corresponding labelled skeletons having six exterior vertices are non-isomorphic.

over $x_1, x_2, \dots, x_n$, the right side becomes

$$\sum_S \sum_\psi \left\{ \prod_{\langle u,v\rangle}{}' \tau_p\big(\psi(u), \psi(v)\big) \right\} \left\{ \sum_x \tau_p(z, x) \right\}^n = \chi(p)^n \sum_S \sum_\psi \prod_{\langle u,v\rangle}{}' \tau_p\big(\psi(u), \psi(v)\big),$$

where z is a typical (but fixed) vertex, the second summation is over all mappings $\psi : I(S) \cup \{0\} \to \mathbb{Z}^d$ satisfying $\psi(0) = x_0$, and the primed product is over all edges $\langle u, v\rangle$ of S as u and v range over the set $I(S) \cup \{0\}$; we have written 0 for the exterior vertex of S with label 0. We now perform the summation over admissible mappings ψ to obtain

$$E_p(|C|^n) \le \chi(p)^n \sum_S \left\{ \sum_y \tau_p(z, y) \right\}^{n-1} = N_{n+1}\chi(p)^{2n-1}, \tag{6.94}$$

where N_{n+1} is the number of labelled skeletons with $n+1$ exterior vertices.

We now compute N_{n+1}. Any labelled skeleton with $n+1$ exterior vertices has $2n-1$ edges. If we remove the exterior vertex labelled n, together with the incident edge and corresponding interior vertex, we obtain a labelled skeleton with n exterior vertices and $2n-3$ edges (see Figure 6.9). Exactly $2n-3$ labelled skeletons with $n+1$ exterior vertices correspond to the resulting labelled skeleton with n exterior vertices, since the $(n+1)$th exterior vertex may be connected into any of its $2n-3$ edges. It is not difficult to convince oneself that these $2n-3$ labelled skeletons are non-isomorphic, and it follows that

$$N_{n+1} = (2n-3)N_n. \tag{6.95}$$

However, $N_3 = 1$, so that

$$N_{n+1} = (2n-3)(2n-5)\dots 5 \cdot 3 \cdot 1 = \frac{(2n-2)!}{2^{n-1}(n-1)!}. \tag{6.96}$$

We combine this with (6.86) and (6.94) to find that

$$(6.97)\quad \begin{aligned} E_p(|C|e^{t|C|}) &= \sum_{n=0}^{\infty} \frac{1}{n!} t^n E_p(|C|^{n+1}) \\ &\le \chi(p)\left\{1 + \sum_{n=1}^{\infty} \frac{1}{n!} t^n N_{n+2}\chi(p)^{2n}\right\} \quad \text{by (6.94)} \\ &= \chi(p)\big(1 - 2t\chi(p)^2\big)^{-1/2} \quad \text{by (6.96)} \end{aligned}$$

for t satisfying $0 \le t < \frac{1}{2}\chi(p)^{-2}$. We now use Markov's inequality (see Grimmett and Stirzaker (1992, p. 285)) to obtain

$$(6.98)\qquad P_p(|C| \ge n) = P_p\big(|C|e^{t|C|} \ge ne^{tn}\big) \le \frac{1}{ne^{tn}} E_p(|C|e^{t|C|})$$

for $t \ge 0$. Applying (6.97), we find that

$$(6.99)\qquad P_p(|C| \ge n) \le \frac{\chi(p)}{ne^{tn}}\big(1 - 2t\chi(p)^2\big)^{-1/2}$$

for values of t satisfying $0 \le t < \frac{1}{2}\chi(p)^{-2}$. We choose

$$t = \frac{1}{2\chi(p)^2} - \frac{1}{2n},$$

noting that $t > 0$ if $n > \chi(p)^2$, to obtain

$$(6.100)\qquad P_p(|C| \ge n) \le (e/n)^{1/2} \exp\big(-\tfrac{1}{2}n/\chi(p)^2\big)$$

in agreement with (6.77). We have proved Theorem (6.75). □

Proof of Theorem (6.78). We follow Kunz and Souillard (1978, p. 91). For ease of description we restrict ourselves to the two-dimensional case, but the same argument is valid whenever $d \ge 2$. Suppose that $0 < p = 1 - q < 1$, and let

$$(6.101)\qquad \pi_n = P_p(|C| = n).$$

We shall show the following lemma.

(6.102) Lemma. *If* $0 < p = 1 - q < 1$, *then* $\pi_k = P_p(|C| = k)$ *satisfies*

$$(6.103)\qquad \frac{1}{m+n}\pi_{m+n} \ge pq^{-2}\frac{1}{m}\pi_m \frac{1}{n}\pi_n \qquad \textit{for all } m, n \ge 1.$$

The limit theorem for subadditive sequences will then yield the result.

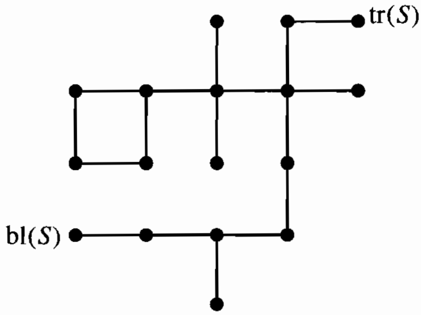


Figure 6.10. A connected subgraph S of $\mathbb{L}^2$, with the 'top right' and 'bottom left' vertices indicated.

Proof of Lemma (6.102). For any connected subgraph S of $\mathbb{L}^2$, we denote by $\mathrm{tr}(S)$ the top vertex on the right side of S. More specifically, $\mathrm{tr}(S) = x$ if $x_1 = \max\{y_1 : y \in S\}$ and $x_2 = \max\{y_2 : (x_1, y_2) \in S\}$. We define $\mathrm{bl}(S)$ similarly as the bottom vertex on the left side of S. See Figure 6.10.

Let π'_k be the probability that the open cluster C at the origin has exactly k vertices and $\mathrm{bl}(C) = 0$. The origin may generally be any of the k vertices of C, and each such possibility has the same probability, so that

$$\pi'_k = \frac{1}{k}\pi_k. \tag{6.104}$$

Let m and n be positive integers. Suppose that σ and τ are connected subgraphs of $\mathbb{L}^2$ with $\mathrm{bl}(\sigma) = \mathrm{bl}(\tau) = 0$, such that σ has m vertices and τ has n vertices. From σ and τ we may construct another subgraph, written $\sigma * \tau$, by translating τ to a new position in $\mathbb{L}^2$ in such a way that $\mathrm{bl}(\tau)$ is translated to the vertex immediately to the right of $\mathrm{tr}(\sigma)$; we add the edge joining $\mathrm{tr}(\sigma)$ to the new position of $\mathrm{bl}(\tau)$ so that the new subgraph $\sigma * \tau$ is connected and has $m + n$ vertices. See Figure 6.11 for an illustration of this construction. It is easy to see that

$$P_p(C = \sigma * \tau) = pq^{-2} P_p(C = \sigma) P_p(C = \tau), \tag{6.105}$$

where the term pq^{-2} arises from the fact that we have declared open an edge which was counted as closed for both σ and τ. Thus two clusters with sizes m and n may be combined to form a cluster with size $m + n$. Furthermore, this composite cluster is unique in the sense that $\sigma * \tau$ may be decomposed in a unique appropriate way into two clusters with sizes m and n. Now

$$\sum_{\sigma} P_p(C = \sigma) = P_p\big(|C| = m,\ \mathrm{bl}(C) = 0\big) = \pi'_m, \tag{6.106}$$

and similarly

$$\sum_{\tau} P_p(C = \tau) = \pi'_n,$$

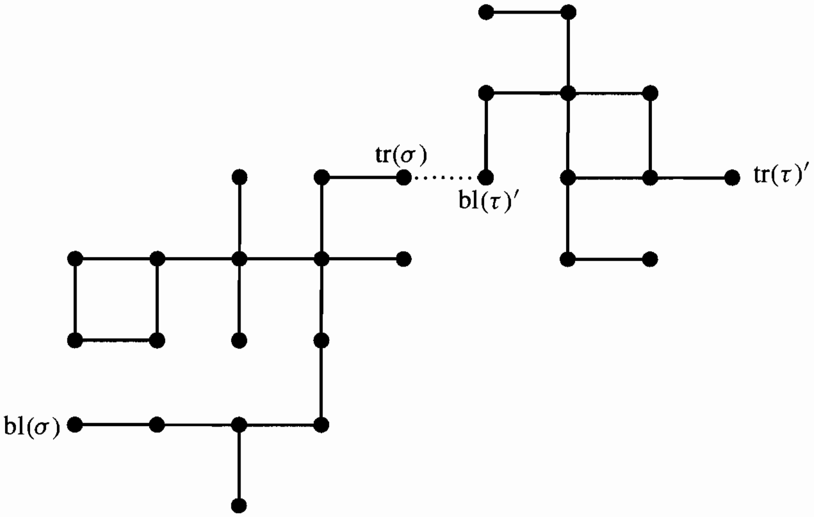


Figure 6.11. The graph σ together with a translation of the graph τ, the bottom left vertex of τ being translated to the immediate right of the top right vertex of σ. The new positions of $\mathrm{bl}(\tau)$ and $\mathrm{tr}(\tau)$ are as indicated by the primed expressions.

where the sums are over all σ and τ with $\mathrm{bl}(\sigma) = \mathrm{bl}(\tau) = 0$ and having m and n vertices, respectively. Therefore,

$$\begin{aligned} pq^{-2}\pi'_m\pi'_n &= \sum_{\sigma,\tau} pq^{-2}P_p(C=\sigma)P_p(C=\tau) \\ &= \sum_{\sigma,\tau} P_p(C=\sigma*\tau) \qquad \text{by (6.105)} \\ &\le P_p\big(|C|=m+n,\ \mathrm{bl}(C)=0\big) = \pi'_{m+n}, \end{aligned}$$

and (6.103) follows from (6.104). □

We have shown that the sequence $(a-\log(\pi_n/n) : n \ge 1)$ is subadditive, where $a = -\log(pq^{-2})$. We apply the subadditive limit theorem (see Appendix II) to find that the limit

$$\zeta(p) = \lim_{n\to\infty}\left\{-\frac{1}{n}\log(\pi_n/n)\right\} \tag{6.107}$$

exists, and

$$a - \log(\pi_n/n) \ge n\zeta(p) \qquad \text{for all } n$$

as required. It is obvious that $p^{n-1}q^{2n+2} \le \pi_n \le 1$, giving that

$$0 \le \zeta(p) \le -\log(pq^2).$$

The strict positivity of $\zeta(p)$ when $0 < p < p_c$ follows from Theorem (6.75). □

6.4 Analyticity of κ and χ

All is peaceful when p is less than p_c. In particular, estimates on the tail of the sizes of open clusters imply the analyticity of macroscopic quantities such as the number of open clusters per vertex and the mean cluster size.

(6.108) Theorem. Analyticity of κ and χ below p_c. *The number $\kappa(p)$ of open clusters per vertex, and the mean size $\chi(p)$ of the open cluster at the origin, are analytic functions of p on the interval* $[0, p_c)$.

Proof. We shall prove that χ is analytic on $[0, p_c)$; the same proof works for κ. Suppose that $p = 1 - q < p_c$. We write $\chi(p)$ in the form (1.20),

$$\chi(p) = \sum_{n=1}^{\infty} n P_p(|C| = n), \tag{6.109}$$

valid since $p < p_c$, where

$$P_p(|C| = n) = \sum_{m,b} a_{nmb} p^m q^b \tag{6.110}$$

as in (4.17). Here, a_{nmb} is the number of animals with n vertices, m edges, and b 'boundary edges'; note that the series in (6.110) is finite.

We follow the calculation of Kesten (1982, p. 250). We recall first that $a_{nmb} = 0$ unless n, m, and b satisfy

$$1 \le b \le 2dn, \qquad n - 1 \le m \le dn \tag{6.111}$$

(see (4.14) and (4.15)), and we assume therefore that these inequalities are satisfied. We may write (6.109) in the form

$$\chi(p) = \sum_n \sum_{m=n-1}^{dn} \sum_{b=1}^{2dn} n a_{nmb} p^m q^b. \tag{6.112}$$

Let z be a complex number, and write

$$K(z) = \sum_n \sum_{m=n-1}^{dn} \sum_{b=1}^{2dn} n a_{nmb} z^m (1 - z)^b \tag{6.113}$$

whenever this sum converges. The function K is a sum of polynomials in z, and thus K is analytic on any disc in which the series in (6.113) is uniformly convergent. We show first that K is analytic on a neighbourhood of the origin of the complex plane. We have from (4.24) that, for fixed n,

$$\sum_{m,b} a_{nmb} \le 7^{dn},$$

and therefore,

$$\text{(6.114)} \qquad \left|\sum_{m,b} n a_{nmb} z^m (1-z)^b\right| \le \sum_{m=n-1}^{dn} \sum_{b=1}^{2dn} n a_{nmb} |z|^m (1+|z|)^{2dn}$$
$$\le n 7^{dn} |z|^{n-1} (1+|z|)^{2dn}$$
$$\le Anc(z)^{n-1}$$

if $|z| < \frac{1}{2}$, where A depends on d only and

$$c(z) = |z|\left\{7(1+|z|)^2\right\}^d.$$

However, $c(z) < 1$ if $|z|$ is sufficiently small, giving from (6.113) and (6.114) that $K(z)$ is uniformly convergent when $|z|$ is sufficiently small; thus K is analytic on a neighbourhood of the origin.

Suppose now that $0 < \alpha < \beta < p_c$. We shall show that K is uniformly convergent on a region of the complex plane containing the real interval $[\alpha, \beta]$ in its interior. Suppose $p \in [\alpha, \beta]$. If $|z - p| < \delta$ and δ is small,

$$\text{(6.115)} \qquad \left|a_{nmb} z^m (1-z)^b\right| \le a_{nmb} p^m q^b \left(\frac{p+\delta}{p}\right)^m \left(\frac{q+\delta}{q}\right)^b$$
$$\le a_{nmb} p^m q^b c(p,\delta)^{dn}$$

by (6.111), where

$$\text{(6.116)} \qquad c(p,\delta) = \frac{p+\delta}{p}\left(\frac{q+\delta}{q}\right)^2.$$

Hence, for fixed n,

$$\sum_{m,b} n|a_{nmb} z^m (1-z)^b| \le nc(p,\delta)^{dn} \sum_{m,b} a_{nmb} p^m q^b$$
$$= nc(p,\delta)^{dn} P_p(|C| = n)$$
$$\le nc(p,\delta)^{dn} P_p(|C| \ge n).$$

The right side here is no larger than

$$nc(p,\delta)^{dn} P_\beta(|C| \geq n) \leq nc(p,\delta)^{dn} e^{-n\lambda(\beta)}$$

for some $\lambda(\beta) > 0$, by Theorem (6.75), giving that

$$\sum_{m,b} n|a_{nmb} z^m (1-z)^b| \leq n\{c(p,\delta)^d e^{-\lambda(\beta)}\}^n.$$

There exists a positive value of δ which is sufficiently small to ensure that

$$c(p,\delta)^d e^{-\lambda(\beta)} < 1-\delta \qquad \text{for } p \in [\alpha, \beta],$$

giving from (6.113) that K converges uniformly on a region of the complex plane containing the real interval $[\alpha, \beta]$, whence K is analytic on $[0, p_c)$ as required.□

6.5 Notes

Section 6.1. Theorem (6.1) was proved by Hammersley (1957a), who did not have the advantage of the BK inequality. Inequality (6.8) appears in van den Berg and Kesten (1985). Theorem (6.10) was proved by Grimmett (1981c); see also Kesten (1987e). Related inequalities for two dimensions appear in Grimmett (1981a). The proof of Theorem (6.14) comes mostly from Grimmett (1983), the principal exception being the proof of the strict monotonicity of φ on $(0, p_c)$. Van den Berg (1983) pointed out that the ad hoc two-dimensional arguments of Grimmett (1983) may be replaced by the more general argument from reliability theory. Chayes and Chayes (1986a, b) have dealt with similar material.

Section 6.2. Excepting the lower bound for $\tau_p(0, e_n)$, Theorem (6.44) is in the folklore of percolation theory. Proposition (6.49) comes directly from Aizenman and Newman (1984); see also Durrett (1985a). For the origins of correlation length, see Chayes, Chayes, and Fröhlich (1985) and Chayes and Chayes (1986a, b). Inequalities such as (6.7) and (6.68) are sometimes named after Simon (1980) and Lieb (1980), who derived a related inequality for the Ising model and certain other systems of mathematical physics; similar inequalities have been used by Hammersley (1957a) and Kesten (1980b). More general forms of such inequalities are at the heart of the theory developed by Aizenman and Newman (1984) and Hara and Slade (1990, 1994) and others concerning critical percolation for large d. See Section 10.3.

The exact polynomial correction for $\tau_p(0, e_n)$ in (6.46) was proved by Campanino, Chayes, and Chayes (1991) for general d (≥ 2). Amongst other things, they established that

$$\tau_p(0, e_n) \sim A(p) \frac{1}{n^{(d-1)/2}} e^{-n/\xi(p)} \qquad \text{as } n \to \infty$$

when $p < p_c$. Here, $A(p)$ is a function of p alone. This completes for percolation a picture proposed by Ornstein and Zernike (1915) and believed to be canonical for a wide variety of disordered physical systems.

Alexander (1990, 1997) has established corrections to the lower bound for $\tau_p(0, x)$ when x lies off the axes of $\mathbb{Z}^d$. Exponential bounds have been used by Alexander, Chayes, and Chayes (1990) and Cerf (1998b) in their verification of the 'Wulff shape' for large finite clusters of supercritical percolation, and by Alexander (1992) in his study of fluctuations for such clusters in two dimensions.

Section 6.3. Theorem (6.75) was proved first by Kesten (1981) by a 'block' approach; previously the exponential decay of the size of C was known only when p is sufficiently close to 0. Our proof is a re-working of the argument of Aizenman and Newman (1984), using the BK inequality. The (near-) supermultiplicativity of $P_p(|C| = n)$ was observed by Kunz and Souillard (1978). Theorem (6.78) and its ramifications may be used (see Grimmett (1985b, 1987a)) to estimate the size of the largest open cluster, longest open path, and longest open circuit of $B(n)$ for large n.

Section 6.4. The analyticity of κ was explored by Kunz and Souillard (1978) when p is sufficiently small. Grimmett (1981b) showed that κ is infinitely differentiable for $0 < p < p_c$ when $d = 2$, and Kesten (1981) proved analyticity for all d. All such results apply to χ also.

Chapter 7

Dynamic and Static Renormalization

7.1 Percolation in Slabs

This chapter and the next are devoted to supercritical percolation, that is, the case $p > p_c$. When $d = 2$, no fundamentally new arguments are needed, since most properties of the supercritical process may be studied via its subcritical dual; see Chapter 11. The picture is however more complicated when $d \geq 3$. The principal target of this chapter is to describe an argument which enables a full study of supercritical percolation when $d \geq 3$.

For any connected graph G, we write $p_c(G)$ for the critical probability of bond percolation on G; this is well defined, by Theorem (2.8). Let $d \geq 3$. We define the (two-dimensional) slab of thickness k by

$$S_k = \mathbb{Z}^2 \times \{0, 1, 2, \dots, k\}^{d-2}, \tag{7.1}$$

having critical probability $p_c(S_k)$. Since $S_k \subseteq S_{k+1} \subseteq \mathbb{Z}^d$ for all k, we have that $p_c(S_k) \geq p_c(S_{k+1}) \geq p_c$. It follows that the decreasing limit

$$p_c^{\text{slab}} = \lim_{k\to\infty} p_c(S_k)$$

exists and satisfies $p_c^{\text{slab}} \geq p_c$.

We say that there is 'percolation in slabs' when $p > p_c^{\text{slab}}$. For values p satisfying $p > p_c^{\text{slab}}$, many facts about the geometry of open clusters may be derived, using the fact that all translates of S_k contain (almost surely) an infinite open cluster. As an example of 'percolation in slabs', we direct the reader to Theorem (8.21), where an exponential decay theorem is proved under this hypothesis. It has been a significant problem to prove that $p_c^{\text{slab}} = p_c$. This and more is proved in the next section, at Theorem (7.2).

Theorem (7.2) was proved by Grimmett and Marstrand (1990), using ideas similar to those of Barsky, Grimmett, and Newman (1991a, b). The proof is achieved via a 'block construction' which appears to be central to a full understanding of supercritical percolation, and to have applications elsewhere. The basic idea is as follows. Suppose that $p > p_c$. We aim to partition $\mathbb{Z}^d$ into blocks having some fixed size M, and to define events associated with such blocks in such a way that three basic things happen. First, we want these events to have probabilities which may be made as close to 1 as required, by a sufficiently large choice of M. Secondly, we wish these events to be independent, or at least 'not too dependent'. And thirdly, we hope that an infinite cluster of good blocks (that is, blocks whose associated events actually occur) will correspond to an infinite open cluster of the percolation process in the original lattice. If these three things can be arranged in a suitable way, then one will argue as follows. Since each block is good with large probability ($1 - \varepsilon$, say, where $\varepsilon = \varepsilon(M)$), one may compare the process of good blocks with a site percolation process on the block lattice having some site-density $f(M, r)$, where r is the range of dependence of the block process. Furthermore, $f(M, r) \to 1$ as $M \to \infty$. Let ρ satisfy $p_c^{\text{site}}(\mathbb{Z}^2) < \rho < 1$. By picking M sufficiently large, one may arrange that the set of good blocks dominates (in the sense of stochastic domination) a site percolation process having density ρ. Since $\rho > p_c^{\text{site}}(\mathbb{Z}^2)$, there exists a.s. an infinite cluster of good blocks within a two-dimensional region of the original lattice. There must then exist a corresponding infinite open cluster within S_N for some large N. This implies that $p > p_c(S_N)$, and the formula $p_c^{\text{slab}} = p_c$ follows.

There are several complications in implementing this programme. The required block argument is sometimes referred to as a 'renormalization' argument. It turns out in the present case that the renormalization is 'dynamic' rather than 'static', in the sense that the events associated with different blocks are defined sequentially in terms of those defined already, rather than being fixed ahead of time.

When correctly formulated, one achieves in this way a method for reducing questions about supercritical percolation for $d \geq 3$ to questions about a supercritical two-dimensional process. In rough terms, one might say that results achievable in two dimensions have counterparts for $d \geq 3$.

We present in Section 7.2 the details of the programme summarized above. It turns out that the methods shed some light upon the question of whether or not $\theta(p_c) = 0$. We explore this in Section 7.3, where results of Barsky, Grimmett, and Newman (1991a, b) concerning percolation in a half-space of $\mathbb{L}^d$ are described.

Whereas renormalization has provided valuable but usually non-rigorous insight to physicists (see Section 9.3), the more complicated constructions of Sections 7.2 and 7.3 are mathematically rigorous. Their dynamic quality referred to above can impede their easy application, and therefore we include in Section 7.4 a type of 'static' renormalization which is validated via the dynamic renormalization, and which can be rather useful in practice.

7.2 Percolation of Blocks

Subject to the hypothesis $\theta(p) > 0$, we wish to gain some understanding over the geometry of the (a.s.) unique open cluster. In particular we shall prove the following theorem, in which $p_c(A)$ denotes the critical value of bond percolation on the subgraph of $\mathbb{Z}^d$ induced by the vertex set A. In this notation, $p_c = p_c(\mathbb{Z}^d)$.

(7.2) Theorem. Limit of slab critical points.

(a) *Let $d \geq 2$, and let F be an infinite connected subset of $\mathbb{Z}^d$ with $p_c(F) < 1$. For each $\eta > 0$ there exists an integer k such that*

$$p_c\big(2kF + B(k)\big) \leq p_c + \eta. \tag{7.3}$$

(b) *If $d \geq 3$, we have that $p_c^{\text{slab}} = p_c$.*

It is easy to see as follows that (a) implies (b). Suppose that $d \geq 3$. Choosing $F = \mathbb{Z}^2 \times \{0\}^{d-2}$, we have that

$$2kF + B(k) = \big\{x \in \mathbb{Z}^d : -k \leq x_j \leq k \text{ for } 3 \leq j \leq d\big\}, \tag{7.4}$$

which is a translate of the slab S_{2k}. Part (a) of the theorem then implies that

$$p_c(S_{2k}) = p_c\big(2kF + B(k)\big) \to p_c \quad \text{as } k \to \infty,$$

whence $p_c^{\text{slab}} = p_c$. It will therefore suffice to prove part (a).

Before proceeding to the formal proof, we sketch the salient features of the block construction necessary to prove the above theorem. This construction may be used directly to obtain further information concerning supercritical percolation.

The main idea involves working with a 'block lattice' each vertex of which represents a large box of $\mathbb{L}^d$, these boxes being disjoint and adjacent. In this block lattice, we declare a vertex to be 'occupied' if there exist certain open paths in and near the corresponding box of $\mathbb{L}^d$. We shall show that, with positive probability, there exists an infinite path of occupied vertices in the block lattice. Furthermore, the corresponding infinite path of blocks of $\mathbb{L}^d$ contains an infinite open path. By choosing sufficiently large boxes, we aim to find such a path within a 'thickening' $2kF + B(k)$ of F, for some sufficiently large k. Thus there is a probabilistic part of the proof, and a geometric part.

There are two main steps in the proof. In the first, we show the existence of certain long finite paths. In the second, we show how to take such finite paths and to build an infinite cluster within a specified subset of the lattice.

The principal parts of the first step are as follows. Suppose that p is such that $\theta(p) > 0$, and that F ($\subseteq \mathbb{Z}^d$) satisfies the conditions of part (a) of the theorem.

1. Let $\varepsilon > 0$. Since $\theta(p) > 0$, there exists m such that
$$P_p\big(B(m) \leftrightarrow \infty\big) > 1 - \varepsilon.$$
This is elementary probability theory.
2. Let $k \geq 1$. We may choose n ($> 2m$) sufficiently large to ensure that, with probability at least $1 - 2\varepsilon$, $B(m)$ is joined to at least k points in $\partial B(n)$.
3. By choosing k sufficiently large, we may ensure that, with probability at least $1 - 3\varepsilon$, $B(m)$ is joined to some point of $\partial B(n)$, which is itself connected to a translate of $B(m)$, lying 'on' the surface $\partial B(n)$ and every edge of which is open.
4. This open translate of $B(m)$, constructed above, may be used as a 'seed' for iterating the above construction. When doing this, we shall need some control over where the seed is placed. It may be shown that, for every face of $\partial B(n)$, $B(m)$ is joined (with large probability) to some point on that face which is itself adjacent to some seed, and indeed to many such points.

Above is the scheme for constructing long finite paths, and we turn to the second step.

5. The above construction is now iterated. At each stage there is a certain (small) probability of failure. In order that there be a strictly positive probability of an infinite sequence of successes, we iterate in two orthogonal directions. With care, one may show that the block construction dominates a certain supercritical site percolation process on F.
6. We wish to deduce that an infinite sequence of successes entails an infinite open path within a region of the form $2kF + B(k)$. There are two difficulties with this. First, since there is not total control of the positions of the seeds, the actual path in $\mathbb{L}^d$ may leave $2kF + B(k)$ for every k. This may be overcome by a process of 'steering', in which, at each stage, we choose a seed in such a position as to compensate for any earlier lateral deviation.
7. A more significant problem is that, while iterating the construction, we carry with us a mixture of 'positive' and 'negative' information (of the form that 'certain paths exist' and 'others do not'). Since some such information is negative, we are not permitted to use the FKG inequality when bounding the probabilities of certain composite events. The practical difficulty is that, although we may achieve an infinite sequence of successes, there will generally be breaks in any corresponding open route to infinity. This difficulty may be overcome by sprinkling down a few more open edges, that is, by working at edge-density $p + \eta$ where $\eta > 0$, rather than at p.

In conclusion, we shall show that, if $\theta(p) > 0$ and $\eta > 0$, there is a k such that there exists (with strictly positive probability) an infinite $(p + \eta)$-open path in the region $2kF + B(k)$. This implies that $p_c(2kF + B(k)) \leq p + \eta$ for large k, as required.

Proof of Theorem (7.2)(a). We follow Grimmett and Marstrand (1990). From now on we suppose that $d = 3$; the construction is simpler when $d = 2$, and is

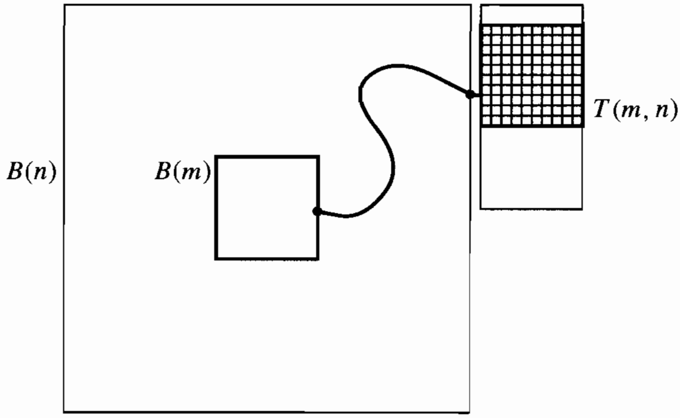


Figure 7.1. An illustration of the event in (7.10). The hatched region is a translate of $B(m)$ all of whose edges are p-open. The central box $B(m)$ is joined by a path to some vertex in $\partial B(n)$, which is in turn connected by a single open edge to a seed lying on the surface of $B(n)$.

similar when $d > 3$. We begin with some notation and two key lemmas. As usual, $B(n) = [-n, n]^3$, and we shall concentrate on a special face of $B(n)$,

$$F(n) = \{x \in \partial B(n) : x_1 = n\}, \tag{7.5}$$

and indeed on a special 'quadrant' of $F(n)$,

$$T(n) = \{x \in \partial B(n) : x_1 = n,\ x_j \geq 0 \text{ for } j \geq 2\}. \tag{7.6}$$

For $m, n \geq 1$, let

$$T(m, n) = \bigcup_{j=1}^{2m+1} \{je_1 + T(n)\} \tag{7.7}$$

where $e_1 = (1, 0, 0)$ as usual.

Let $m \geq 1$. We call a box $x + B(m)$ a *seed* if every edge in $x + B(m)$ is open. We now set

$$K(m, n) = \Big\{x \in T(n) : \langle x, x + e_1\rangle \text{ is open, and } x + e_1 \text{ lies in some seed lying within } T(m, n)\Big\}. \tag{7.8}$$

The random set $K(m, n)$ is necessarily empty if $n < 2m$.

(7.9) Lemma. *If $\theta(p) > 0$ and $\eta > 0$, there exist integers $m = m(p, \eta)$ and $n = n(p, \eta)$ such that $2m < n$ and*

$$P_p\big(B(m) \leftrightarrow K(m, n) \text{ in } B(n)\big) > 1 - \eta. \tag{7.10}$$

The event in (7.10) is illustrated in Figure 7.1.

Proof. Since $\theta(p) > 0$, there exists a.s. an infinite open cluster, whence

$$P_p\big(B(m) \leftrightarrow \infty\big) \to 1 \quad \text{as} \quad m \to \infty.$$

We pick m such that

$$P_p\big(B(m) \leftrightarrow \infty\big) > 1 - (\tfrac{1}{3}\eta)^{24}, \tag{7.11}$$

for a reason which will become clear later.

For $n > m$, let $V(n) = \{x \in T(n) : x \leftrightarrow B(m) \text{ in } B(n)\}$. Pick M such that

$$pP_p\big(B(m) \text{ is a seed}\big) > 1 - (\tfrac{1}{2}\eta)^{1/M}. \tag{7.12}$$

We shall assume for simplicity that $2m + 1$ divides $n + 1$, and we partition $T(n)$ into disjoint squares with side-length $2m$. If $|V(n)| \geq (2m+1)^2 M$ then $B(m)$ is joined in $B(n)$ to at least M of these squares. Therefore, by (7.12),

(7.13)
$$\begin{aligned} &P_p\big(B(m) \leftrightarrow K(m,n) \text{ in } B(n)\big) \\ &\qquad \geq \Big\{1 - \big[1 - pP_p\big(B(m) \text{ is a seed}\big)\big]^M\Big\} P_p\big(|V(n)| \geq (2m+1)^2 M\big) \\ &\qquad \geq (1 - \tfrac{1}{2}\eta) P_p\big(|V(n)| \geq (2m+1)^2 M\big). \end{aligned}$$

We now bound the last probability from below. Using the symmetries of $\mathbb{L}^3$ obtained by reflections in hyperplanes, we see that the face $F(n)$ comprises four copies of $T(n)$. Now $\partial B(n)$ has six faces, and therefore 24 copies of $T(n)$. By symmetry and the FKG inequality,

$$P_p\big(|U(n)| < 24(2m+1)^2 M\big) \geq P_p\big(|V(n)| < (2m+1)^2 M\big)^{24} \tag{7.14}$$

where

$$U(n) = \big\{x \in \partial B(n) : x \leftrightarrow B(m) \text{ in } B(n)\big\}.$$

Now, with $l = 24(2m+1)^2 M$,

$$P_p\big(|U(n)| < l\big) \leq P_p\big(|U(n)| < l,\ B(m) \leftrightarrow \infty\big) + P_p\big(B(m) \not\leftrightarrow \infty\big), \tag{7.15}$$

and

(7.16)
$$\begin{aligned} P_p\big(|U(n)| < l,\ B(m) \leftrightarrow \infty\big) &\leq P_p\big(1 \leq |U(n)| < l\big) \\ &\leq (1-p)^{-3l} P_p\big(U(n+1) = \varnothing,\ U(n) \neq \varnothing\big) \\ &\to 0 \qquad \text{as } n \to \infty. \end{aligned}$$

(We have used the fact that $U(n+1) = \varnothing$ if every edge exiting $\partial B(n)$ from $U(n)$ is closed.)

By (7.14)–(7.16) and (7.11),

$$P_p\big(|V(n)| < (2m+1)^2 M\big) \le P_p\big(|U(n)| < l\big)^{1/24} \le \big(a_n + (\tfrac{1}{3}\eta)^{24}\big)^{1/24}$$

where $a_n \to 0$ as $n \to \infty$. We pick n such that

$$P_p\big(|V(n)| < (2m+1)^2 M\big) \le \tfrac{1}{2}\eta,$$

and the claim of the lemma follows by (7.13). □

Having constructed open paths from $B(m)$ to $K(m,n)$, we shall need to repeat the construction, beginning now at an appropriate seed in $T(m,n)$. This is problematic, since we have discovered a mixture of information, some of it negative, about the immediate environs of such seeds. In order to overcome the effect of such negative information, we shall work at edge-density $p+\delta$ rather than p. In preparation, let $(X(e) : e \in \mathbb{E}^d)$ be independent random variables having the uniform distribution on $[0,1]$, and let $\eta_p(e)$ be the indicator function that $X(e) < p$; recall equation (1.4). As before, we say that e is *p-open* if $X(e) < p$ and *p-closed* otherwise, and we denote by P the appropriate probability measure.

For any subset V of $\mathbb{Z}^3$, we define its *exterior vertex boundary* $\Delta_{\mathrm{v}} V$ and *edge boundary* ΔV by

$$\Delta_{\mathrm{v}} V = \big\{x \in \mathbb{Z}^3 : x \notin V,\ x \sim y \text{ for some } y \in V\big\},$$
$$\Delta V = \big\{\langle x, y\rangle : x \in V,\ y \in \Delta_{\mathrm{v}} V,\ x \sim y\big\}.$$

We write $\mathbb{E}_V$ for the set of all edges of $\mathbb{L}^3$ joining pairs of vertices in V.

We shall make repeated use of the following lemma, which is illustrated in Figure 7.2.

(7.17) Lemma. *If $\theta(p) > 0$ and $\varepsilon, \delta > 0$, there exist integers $m = m(p,\varepsilon,\delta)$ and $n = n(p,\varepsilon,\delta)$ such that $2m < n$ and with the following property. Let R be such that $B(m) \subseteq R \subseteq B(n)$ and $(R \cup \Delta_{\mathrm{v}} R) \cap T(n) = \varnothing$, and let $\beta : \Delta R \cap \mathbb{E}_{B(n)} \to [0, 1-\delta]$. Define the events*

$G = \{$there exists a path joining R to $K(m,n)$, this path being p-open outside ΔR and $(\beta(e)+\delta)$-open at its unique edge e lying in $\Delta R\}$,

$H = \{e$ is $\beta(e)$-closed for all $e \in \Delta R \cap \mathbb{E}_{B(n)}\}$.

Then $P(G \mid H) > 1 - \varepsilon$.

Proof. Assume that $\theta(p) > 0$, and let $\varepsilon, \delta > 0$. Pick an integer t so large that

$$(1-\delta)^t < \tfrac{1}{2}\varepsilon \qquad (7.18)$$

and then choose $\eta\ (>0)$ such that

$$\eta < \tfrac{1}{2}\varepsilon(1-p)^t. \qquad (7.19)$$

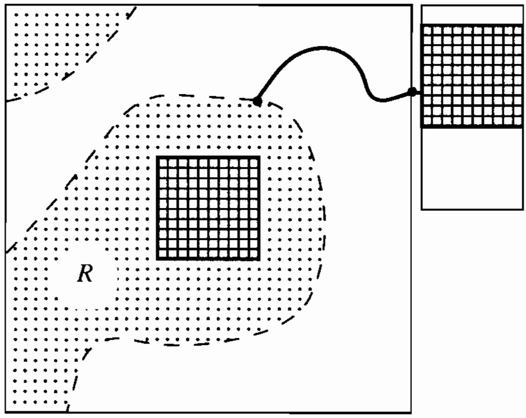


Figure 7.2. An illustration of Lemma (7.17). The hatched regions are translates of $B(m)$ all of whose edges are p-open. The central box $B(m)$ lies within some (dotted) region R. Some vertex in R is joined by a path to some vertex in $\partial B(n)$, which is in turn connected by a single open edge to a seed lying on the surface of $B(n)$.

We apply Lemma (7.9) with this value of η, thereby obtaining integers m, n such that $2m < n$ and

$$(7.20) \qquad P_p\big(B(m) \leftrightarrow K(m,n) \text{ in } B(n)\big) > 1 - \eta.$$

Let R and β satisfy the hypotheses of the lemma. Since any path from $B(m)$ to $K(m,n)$ contains a path from ∂R to $K(m,n)$, we have that

$$(7.21) \qquad P_p\big(\partial R \leftrightarrow K(m,n) \text{ in } B(n)\big) > 1 - \eta.$$

Let $K \subseteq T(n)$, and let $U(K)$ be the set of edges $\langle x, y\rangle$ of $B(n)$ such that

(i) $x \in R$, $y \notin R$; and

(ii) there is an open path joining y to K, using no edges of $\mathbb{E}_R \cup \Delta R$.

We wish to show that $U(K)$ must be large if $P_p(\partial R \leftrightarrow K \text{ in } B(n))$ is large. The argument centres on the fact that every path from ∂R to K passes through $U(K)$; if $U(K)$ is 'small' then there is substantial uncertainty for the occurrence of the event $\{\partial R \leftrightarrow K \text{ in } B(n)\}$, implying that this event cannot have probability near 1. More rigorously,

$$(7.22) \qquad \begin{aligned} P_p\big(\partial R \not\leftrightarrow K \text{ in } B(n)\big) &= P_p\big(\text{all edges in } U(K) \text{ are closed}\big) \\ &\ge (1-p)^t P_p\big(|U(K)| \le t\big). \end{aligned}$$

We may apply this with $K = K(m,n)$, since $K(m,n)$ is defined on the set of edges exterior to $B(n)$. Therefore, by (7.21), (7.22), and (7.19),

(7.23)

$$\begin{aligned} P_p\big(|U(K(m,n))| > t\big) &\ge 1 - (1-p)^{-t} P_p\big(\partial R \not\leftrightarrow K(m,n) \text{ in } B(n)\big) \\ &\ge 1 - (1-p)^{-t}\eta > 1 - \tfrac{1}{2}\varepsilon. \end{aligned}$$

We now couple together the percolation processes with different values of p on the same probability space, as described in Section 1.3 and just prior to the statement of Lemma (7.17). We borrow the notation and results derived above by specializing to the p-open edges. Conditional on the set $U = U(K(m,n))$, the values of $X(e)$, for $e \in U$, are independent and uniform on $[0,1]$. Therefore,

$$P\Big(\text{every } e \text{ in } U \text{ is } (\beta(e)+\delta)\text{-closed}, |U| > t \,\Big|\, H\Big) \le (1-\delta)^t,$$

whence, using (7.18) and (7.23),

$$\begin{aligned} P\Big(\text{some } e \text{ in } U \text{ is } (\beta(e)+\delta)\text{-open} \,\Big|\, H\Big) &\ge P(|U| > t \mid H) - (1-\delta)^t \\ &= P_p(|U| > t) - (1-\delta)^t \\ &\ge (1 - \tfrac{1}{2}\varepsilon) - \tfrac{1}{2}\varepsilon, \end{aligned}$$

and the lemma is proved. □

This completes the two key geometrical lemmas. In moving to the second part of the proof, we shall require a method for comparison of a 'dependent' process and a site percolation process. The argument required at this stage is as follows.

Let F be an infinite connected subset of $\mathbb{L}^d$. We have from Theorem (1.33) that $p_c(F) < 1$ if and only if $p_c^{\text{site}}(F) < 1$, and we assume henceforth that both inequalities hold. Let $\{Z(x) : x \in F\}$ be random variables taking values in $\{0,1\}$. We construct a connected subset of F in the following recursive manner (cf. the proof of Theorem (1.33)). Let $f_1, f_2, \dots$ be a fixed ordering of the edges of the graph induced by F. Let $x_1 \in F$, and define the ordered pair $S_1 = (A_1, B_1)$ of subsets of F by

$$S_1 = \begin{cases} (\{x_1\}, \varnothing) & \text{if } Z(x_1) = 1, \\ (\varnothing, \{x_1\}) & \text{if } Z(x_1) = 0. \end{cases}$$

Having defined $S_1, S_2, \dots, S_t = (A_t, B_t)$, for $t \ge 1$, we define S_{t+1} as follows. Let f be the earliest edge in the fixed ordering $f_1, f_2, \dots$ with the property that one endvertex lies in A_t and the other endvertex, x_{t+1} say, lies outside $A_t \cup B_t$. We declare

$$S_{t+1} = \begin{cases} \big(A_t \cup \{x_{t+1}\}, B_t\big) & \text{if } Z(x_{t+1}) = 1, \\ \big(A_t, B_t \cup \{x_{t+1}\}\big) & \text{if } Z(x_{t+1}) = 0. \end{cases}$$

If no such edge f exists, we declare $S_{t+1} = S_t$. The sets A_t, B_t are non-decreasing, and we set

$$A_\infty = \lim_{t\to\infty} A_t, \quad B_\infty = \lim_{t\to\infty} B_t.$$

We think of A_∞ as the 'occupied cluster' at x_1, and B_∞ as its external vertex boundary.

(7.24) Lemma. *Suppose there exists a constant* γ *such that* $\gamma > p_c^{\text{site}}(F)$ *and*

$$(7.25) \qquad P\big(Z(x_{t+1}) = 1 \mid S_1, S_2, \dots, S_t\big) \ge \gamma \quad \textit{for all } t.$$

Then $P(|A_\infty| = \infty) > 0$.

We omit a formal proof of this lemma, which may be found in Grimmett and Marstrand (1990). Informally, (7.25) states that, uniformly in the past history, the chance of extending A_t exceeds the critical value of a supercritical site percolation process on F. Therefore A_∞ dominates stochastically the open cluster at x_1 of a supercritical site percolation process. The latter cluster is infinite with strictly positive probability, whence $P(|A_\infty| = \infty) > 0$. The method of 'stochastic domination' is often useful in studying percolation and other disordered systems; see, for example, the proof of Theorem (1.33). A more systematic account of stochastic domination is given in Section 7.4, particularly the discussion leading to Theorem (7.65).

Having established the three basic lemmas, we turn to the construction itself. Recall the notation and hypotheses of Theorem (7.2)(a). Let η be small and strictly positive, and choose

$$(7.26) \qquad p = p_c + \tfrac{1}{2}\eta, \quad \delta = \tfrac{1}{12}\eta, \quad \varepsilon = \tfrac{1}{24}\big(1 - p_c^{\text{site}}(F)\big).$$

Recall that $p_c^{\text{site}}(F) < 1$ by the discussion above. Since $p > p_c$, we have that $\theta(p) > 0$, and we apply Lemma (7.17) with the above ε, δ to find corresponding integers m, n. We define $N = m + n + 1$, and we shall define a process on the blocks of $\mathbb{Z}^3$ having side-length $2N$.

Consider the set $\{4Nx : x \in \mathbb{Z}^d\}$ of vertices, and the associated boxes

$$B_x = B_x(N) = 4Nx + B(N) \quad \text{for } x \in \mathbb{Z}^d;$$

these boxes we call *site-boxes*. A pair B_x, B_y of site-boxes is deemed *adjacent* if x and y are adjacent in $\mathbb{L}^d$. Adjacent site-boxes are linked by *bond-boxes*, that is, boxes $Nz + B(N)$ for $z \in \mathbb{Z}^d$ exactly one component of which is not divisible by 4. If this exceptional component of z is even, the box $Nz + B(N)$ is called a *half-way* box. See Figure 7.3.

We shall examine site-boxes one by one, declaring each to be 'occupied' or 'unoccupied' according to the existence (or not) of certain open paths. Two properties of this construction will emerge.

(a) For each site-box considered, the probability that it is occupied exceeds the critical probability of a certain site percolation process. This property will imply later in the proof that, with strictly positive probability, there exists an infinite occupied path of site-boxes.

(b) The existence of this infinite occupied path necessarily entails an infinite open path of $\mathbb{L}^d$ lying within some restricted region.

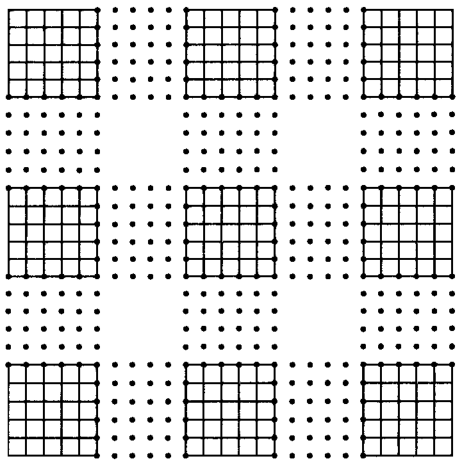


Figure 7.3. The hatched squares are site-boxes, and the dotted squares are half-way boxes. Each box has side-length $2N$.

The site-boxes will be examined in sequence, the order of this sequence being random, and depending on the past history of the process. Thus, the renormalization is 'dynamic' rather than 'static'.

As above, let F be an infinite connected subset of $\mathbb{Z}^d$; we shall assume for convenience that F contains the origin 0 (otherwise, translate F accordingly). As above, let $f_1, f_2, \ldots$ be a fixed ordering of the edges joining vertices in F. We shall examine the site-boxes B_x, for $x \in F$, and determine their states. This we do according to the algorithm sketched before Lemma (7.24) for appropriate random variables $Z(x)$ to be described next.

We begin at the origin, with the site-box $B_0 = B(N)$. Once we have explained what is involved in determining the state of B_0, most of the work will have been done. (The event $\{B_0$ is occupied$\}$ is sketched in Figure 7.5.)

We note that $B(m) \subseteq B(N)$, and we say that 'the first step is successful' if every edge in $B(m)$ is p-open, which is to say that $B(m)$ is a 'seed'. (Recall that p and other parameters are given in (7.26).) At this stage we write E_1 for the set of edges of $B(m)$.

In the following sequential algorithm, we shall construct an increasing sequence $E_1, E_2, \ldots$ of edge-sets. At each stage k, we shall acquire information about the values of $X(e)$ for certain $e \in \mathbb{E}^3$ (here, the $X(e)$ are independent uniform $[0, 1]$-valued random variables, as usual). This information we shall record in the form 'each e is $\beta_k(e)$-closed and $\gamma_k(e)$-open' for suitable functions $\beta_k, \gamma_k : \mathbb{E}^3 \to [0, 1]$ satisfying

$$(7.27) \qquad \beta_k(e) \le \beta_{k+1}(e), \quad \gamma_k(e) \ge \gamma_{k+1}(e), \quad \text{for all } e \in \mathbb{E}^3.$$

Having constructed E_1, above, we set

$$\beta_1(e) = 0 \quad \text{for all } e \in \mathbb{E}^3, \tag{7.28}$$

$$\gamma_1(e) = \begin{cases} p & \text{if } e \in E_1, \\ 1 & \text{otherwise.} \end{cases} \tag{7.29}$$

Since we are working with edge-sets E_j rather than with vertex-sets, it will be useful to have some corresponding notation. Two edges e, f are called *adjacent*, written $e \approx f$, if they have exactly one common endvertex. This adjacency relation defines a graph. Paths in this graph are said to be α-open if $X(e) < \alpha$ for all e lying in the path. The edge boundary ΔE of an edge-set E is the set of all edges $f \in \mathbb{E}^3 \setminus E$ such that $f \approx e$ for some $e \in E$. As before, for $A \subseteq \mathbb{Z}^d$, we write $\mathbb{E}_A$ for the set of all edges having both endvertices in A.

For $j = 1, 2, 3$ and $\sigma = \pm$, let e_j be the jth unit vector of the lattice, and let L_j^σ be an automorphism of $\mathbb{L}^3$ which preserves the origin and which maps $e_1 = (1, 0, 0)$ onto σe_j; we insist that L_1^+ is the identity. We now define E_2 as follows. Consider the set of all paths π lying within the region

$$Z = B(n) \cup \left\{ \bigcup_{\substack{1 \le j \le 3 \\ \sigma = \pm}} L_j^\sigma\big(T(m, n)\big) \right\}$$

such that

(a) the first edge f of π lies in ΔE_1 and is $(\beta_1(f) + \delta)$-open; and

(b) all other edges lie outside $E_1 \cup \Delta E_1$ and are p-open.

We define $E_2 = E_1 \cup F_1$ where F_1 is the set of all edges in the union of such paths π. We say that 'the second step is successful' if, for each $j = 1, 2, 3$ and $\sigma = \pm$, there exists an edge in E_2 having an endvertex in $K_j^\sigma(m, n)$, where

$$K_j^\sigma(m, n) = \Big\{ z \in L_j^\sigma\big(T(n)\big) : \langle z, z + \sigma e_j \rangle \text{ is } p\text{-open, and } z + \sigma e_j \text{ lies in some seed lying within } L_j^\sigma\big(T(m, n)\big) \Big\}.$$

The corresponding event is illustrated in Figure 7.4.

Next we estimate the probability that the second step is successful, conditional on the first step being successful. Let G be the event that there exists a path in $B(n) \setminus B(m)$ from $\partial B(m)$ to $K(m, n)$, every edge e of which is p-open off ΔE_1 and whose unique edge f in ΔE_1 is $(\beta_1(f) + \delta)$-open. We write G_j^σ for the corresponding event with $K(m, n)$ replaced by $L_j^\sigma(K(m, n))$. We now apply Lemma (7.17) with $R = B(m)$ and $\beta = \beta_1$ to find that

$$P\big(G_j^\sigma \,\big|\, B(m) \text{ is a seed}\big) > 1 - \varepsilon \quad \text{for } j = 1, 2, 3,\ \sigma = \pm.$$

Therefore,

$$P\big(G_j^\sigma \text{ occurs for all } j, \sigma \,\big|\, B(m) \text{ is a seed}\big) > 1 - 6\varepsilon, \tag{7.30}$$

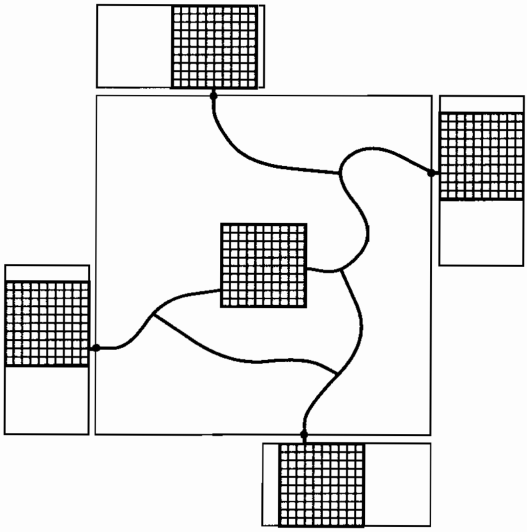


Figure 7.4. An illustration of the first two steps in the construction of the event {0 is occupied}, when these steps are successful. Each hatched square is a seed.

so that the second step is successful with conditional probability at least $1 - 6\varepsilon$.

If the second step is successful, we update the β, γ functions accordingly, setting

$$\beta_2(e) = \begin{cases} \beta_1(e) & \text{if } e \notin \mathbb{E}_Z, \\ \beta_1(e) + \delta & \text{if } e \in \Delta E_1 \backslash E_2, \\ p & \text{if } e \in (\Delta E_2 \backslash \Delta E_1) \cap \mathbb{E}_Z, \\ 0 & \text{otherwise}, \end{cases} \tag{7.31}$$

$$\gamma_2(e) = \begin{cases} \gamma_1(e) & \text{if } e \in E_1, \\ \beta_1(e) + \delta & \text{if } e \in \Delta E_1 \cap E_2, \\ p & \text{if } e \in E_2 \backslash (E_1 \cup \Delta E_1), \\ 1 & \text{otherwise}. \end{cases} \tag{7.32}$$

Suppose that the first two steps have been successful. We next aim to link the appropriate seeds in each $L_j^\sigma(T(m,n))$ to a new seed lying in the bond-box $2\sigma N e_j + B(N)$, that is, the half-way box reached by exiting the origin in the direction σe_j. If we succeed with each of the six such extensions, then we terminate this stage of the process, and declare the vertex 0 of the renormalized lattice to be *occupied*; such success constitutes the definition of the term 'occupied' when applied to 0. See Figure 7.5.

We do not present all the details of this part of the construction, since they are very similar to those already described. Instead we concentrate on describing the basic strategy when $j = 1$ and $\sigma = +$, and on discussing any novel aspects of

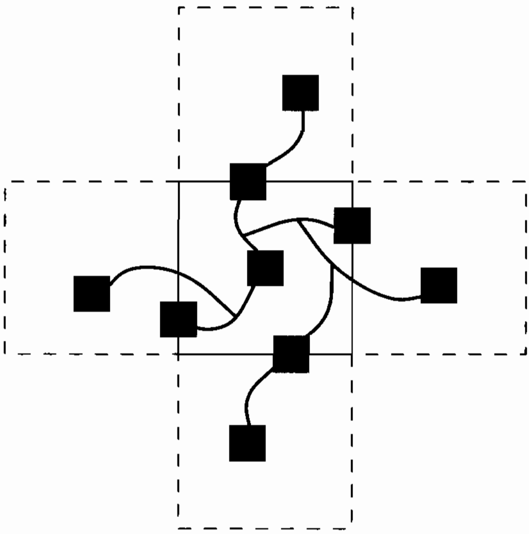


Figure 7.5. An illustration of the event {0 is occupied}. Each black square is a seed.

the construction. First, let $Z_2 = b_2 + B(m)$ be the earliest seed (in some ordering of all copies of $B(m)$) all of whose edges lie in $E_2 \cap \mathbb{E}_{T(m,n)}$. We now try to extend E_2 to include a seed lying within the bond-box $2Ne_1 + B(N)$. Clearly $Z_2 \subseteq Ne_1 + B(N)$. In performing this extension, we encounter a 'steering' problem. It happens (by construction) that all coordinates of b_2 are positive, implying that $b_2 + T(m,n)$ is not a subset of $2Ne_1 + B(N)$. We therefore replace $b_2 + T(m,n)$ by $b_2 + T^*(m,n)$ where $T^*(m,n)$ is given as follows. Instead of working with the 'quadrant' $T(n)$ of the face $F(n)$, we use the set

$$T^*(n) = \{x \in \partial B(n) : x_1 = n,\ x_j \le 0 \text{ for } j = 2, 3\}.$$

We then define

$$T^*(m,n) = \bigcup_{j=1}^{2m+1} \{je_1 + T^*(n)\},$$

and obtain that $b_2 + T^*(m,n) \subseteq 2Ne_1 + B(N)$. We now consider the set of all paths π lying within the region

$$Z_2' = b_2 + \{B(n) \cup T^*(m,n)\}$$

such that:

(a) the first edge f of π lies in ΔE_2 and is $(\beta_2(f) + \delta)$-open; and

(b) all other edges lie outside $E_2 \cup \Delta E_2$ and are p-open.

We set $E_3 = E_2 \cup F_2$ where F_2 is the set of all edges lying in the union of such paths. We call this step successful if E_3 contains an edge having an endvertex in

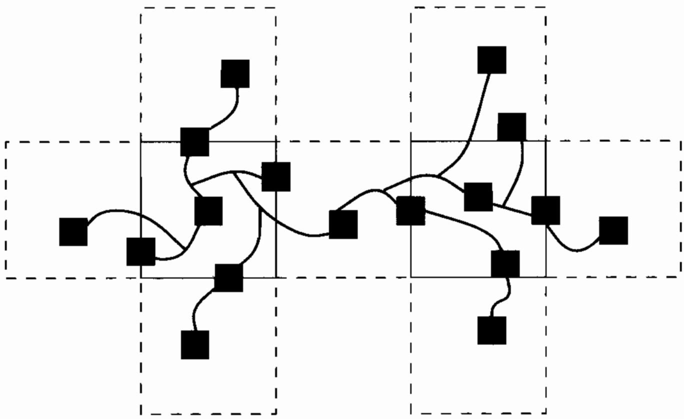


Figure 7.6. Two adjacent site-boxes both of which are occupied. The construction began with the left site-box B_0 and has been extended to the right site-box B_{e_1}. The black squares are seeds, as before.

the set

$$b_2+K^*(m,n) = \Big\{z \in b_2+T^*(m,n) : \langle z, z+e_1\rangle \text{ is } p\text{-open}, z+e_1 \text{ lies in some seed contained in } b_2+T^*(m,n)\Big\}.$$

Using Lemma (7.17), the (conditional) probability that this step is successful exceeds $1-\varepsilon$.

We perform similar extensions in each of the other five directions exiting B_0. If all are successful, we declare 0 to be *occupied*. Combining the above estimates of success, we find that

$$(7.33)\quad P\big(0 \text{ is occupied} \,\big|\, B(m) \text{ is a seed}\big) > (1-6\varepsilon)(1-\varepsilon)^6 > 1-12\varepsilon = \tfrac{1}{2}\big(1+p_c^{\text{site}}(F)\big)$$

by (7.26).

If 0 is not occupied, we end the construction. If 0 is occupied, this has been achieved after the definition of a set E_8 of edges. The corresponding functions β_8, γ_8 are such that

$$(7.34)\qquad \beta_8(e) \le \gamma_8(e) \le p+6\delta \quad \text{for } e \in E_8;$$

this follows since no edge lies in more than 7 of the translates of $B(n)$ used in the repeated application of Lemma (7.17). Therefore, every edge of E_8 is $(p+\eta)$-open, since $\delta = \frac{1}{12}\eta$ (see (7.26)).

The basic idea has been described, and we now proceed similarly. Assume 0 is occupied, and find the earliest edge f_r induced by F and incident with the origin;

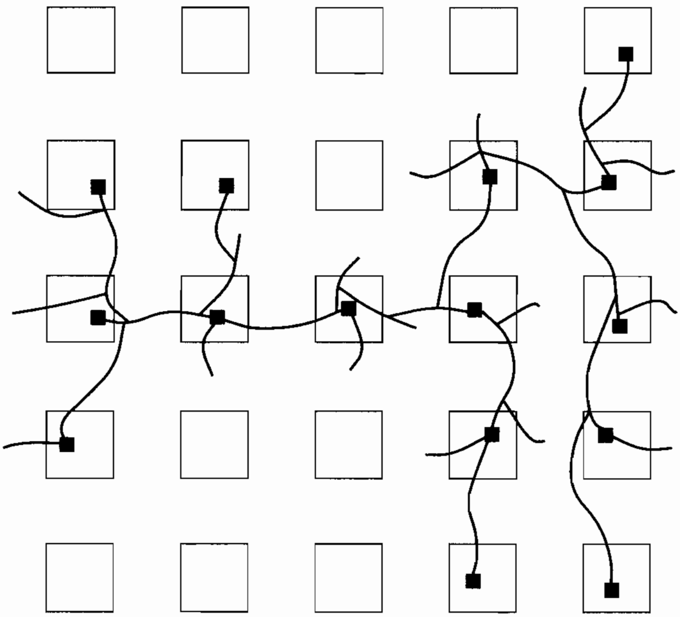


Figure 7.7. The central seed is $B(m)$, and the connections represent $(p+\eta)$-open paths joining seeds within the site-boxes.

we may assume for the sake of simplicity that $f_r = \langle 0, e_1 \rangle$. We now attempt to link the seed $b_3 + B(m)$, found as above inside the half-way box $2Ne_1 + B(N)$, to a seed inside the site-box $4Ne_1 + B(N)$. This is done in two steps of the earlier kind. Having found a suitable seed inside the new site-box $4Ne_1 + B(N)$, we attempt to branch-out in the other 5 directions from this site-box. If we succeed in finding seeds in each of the corresponding half-way boxes, we declare the vertex e_1 of the renormalized lattice to be occupied. As before, the (conditional) probability that e_1 is occupied is at least $\frac{1}{2}(1 + p_c^{\text{site}}(F))$, and every edge in the ensuing construction is $(p + \eta)$-open. See Figure 7.6.

Two details arise at this and subsequent stages, each associated with 'steering'. First, if $b_3 = (\alpha_1, \alpha_2, \alpha_3)$ we concentrate on the quadrant $T_\alpha(n)$ of $\partial B(n)$ defined as the set of $x \in \partial B(n)$ for which $x_j \alpha_j \leq 0$ for $j = 2, 3$ (so that x_j has the opposite sign to α_j). Having found such a $T_\alpha(n)$, we define

$$T_\alpha(m, n) = \bigcup_{j=1}^{2m+1} \left\{ je_1 + T_\alpha(n) \right\},$$

and we look for paths from $b_3 + B(m)$ to $b_3 + T_\alpha(m, n)$. This mechanism guarantees that any variation in b_3 from the first coordinate axis is (at least partly) compensated for at the next step.

A further detail arises when branching out from the seed $b^* + B(m)$ reached inside $4Ne_1 + B(N)$. In finding seeds lying in the new half-way boxes abutting

$4Ne_1 + B(N)$, we 'steer away from the inlet branch', by examining seeds lying on the surface of $b^* + B(n)$ with the property that the first coordinates of their vertices are not less than that of b^*. This process guarantees that these seeds have not been examined previously.

We now continue to apply the algorithm presented before Lemma (7.24). At each stage, the chance of success exceeds $\gamma = \frac{1}{2}(1 + p_c^{\text{site}}(F))$. Since $\gamma > p_c^{\text{site}}(F)$, we have from Lemma (7.24) that there is a strictly positive probability that the ultimate set of occupied vertices of F (that is, renormalized blocks of $\mathbb{L}^3$) is infinite. Now, on this event, there must exist an infinite $(p + \eta)$-open path of $\mathbb{L}^3$ within the enlarged set $4NF + B(2N)$. See Figure 7.7. We deduce that

$$p_c + \eta \geq p_c\big(4NF + B(2N)\big),$$

as required for Theorem (7.2). The proof is complete. □

7.3 Percolation in Half-Spaces

The method of the last section comes close to proving that $\theta(p_c) = 0$ for general d, as the following explains. Suppose $\theta(p) > 0$ and $\eta > 0$. There is effectively defined in Section 7.2 an event A living in a finite box B such that:

(a) $P_p(A) > 1 - \varepsilon$, for some prescribed small $\varepsilon > 0$;

(b) the fact (a) implies that $\theta(p + \eta) > 0$.

Suppose that we could prove this with $\eta = 0$, and assume for the moment that $\theta(p_c) > 0$. Then $P_{p_c}(A) > 1 - \varepsilon$. Now, $P_p(A)$ is a finite polynomial in p, since B is finite. Therefore $P_p(A)$ is a continuous function of p, whence there exists $p' < p_c$ such that $P_{p'}(A) > 1 - \varepsilon$. It follows by (b), with $\eta = 0$, that $\theta(p') > 0$. This conclusion would contradict the definition of p_c, whence we deduce that $\theta(p_c) = 0$.

It is vital for the construction that η be assumed strictly positive, since we need on occasion to 'spend a little extra money' in order to compensate for negative information acquired earlier in the construction. In a slightly different setting, no extra money is required.

Let $d \geq 2$, and let $\mathbb{H} = \mathbb{Z}^{d-1} \times \mathbb{Z}_+$ where $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$. We call $\mathbb{H}$ a *half-space*, and we write

$$\theta_{\mathbb{H}}(p) = P_p(0 \leftrightarrow \infty \text{ in } \mathbb{H})$$

for the associated percolation probability, and

$$p_c(\mathbb{H}) = \sup\{p : \theta_{\mathbb{H}}(p) = 0\}$$

for the critical probability of $\mathbb{H}$. It follows by taking $F = \mathbb{H}$ in Theorem (7.2) that $p_c(\mathbb{H}) = p_c$.

(7.35) Theorem. *Let $d \geq 2$. We have that $\theta_{\mathbb{H}}(p_c) = 0$.*

Such a conclusion for half-spaces has a striking implication for the conjecture that $\theta(p_c) = 0$, namely the following. Suppose $\theta(p_c) > 0$; when $p = p_c$, there exists a.s. a unique infinite open cluster in $\mathbb{Z}^d$, which is a.s. partitioned into (only) finite clusters by *any* division of $\mathbb{Z}^d$ into two half-spaces.

Proof. We follow Barsky, Grimmett, and Newman (1991a, b), making use of a block construction which differs from that of the last section in two major regards. First, since $\mathbb{H}$ lacks some symmetry, we shall have to work with blocks which are not cubical. Secondly, we are able to work at the same value of p throughout; there is no necessity to augment the edge-density. The details of this construction will be given under the assumption that $d = 3$; the proof is simpler when $d = 2$, and a similar argument is valid when $d > 3$.

Our basic building blocks are given as follows. Let L, H be positive integers and define the *brick* $B(L, H)$ by

$$B(L, H) = [-L, L]^2 \times [0, H].$$

(Throughout this proof we shall interpret intervals of the form $[-a, b]$ as intervals of integers.) Note that $B(L, H)$ has 'length' $2L$ and 'height' H. This brick has an underside U, a top T, and sides S, given by

$$\begin{aligned} U = U(L, H) &= [-L, L]^2 \times \{0\}, \\ T = T(L, H) &= [-L, L]^2 \times \{H\}, \\ S = S(L, H) &= \big\{x \in B(L, H) : |x_j| = L \text{ for some } j \in \{1, 2\}\big\}, \end{aligned}$$

The top T may be divided into 4 congruent 'subfacets'

$$\begin{aligned} T_1 &= [0, L]^2 \times \{H\}, & T_2 &= [0, L] \times [-L, 0] \times \{H\}, \\ T_3 &= [-L, 0]^2 \times \{H\}, & T_4 &= [-L, 0] \times [0, L] \times \{H\}. \end{aligned}$$

Similarly, the set S may be divided into 4 'facets', each of which may be divided into 2 'subfacets' congruent to $\{L\} \times [0, L] \times [0, H]$; we denote the 4×2 corresponding subregions of S as $S_1, S_2, \ldots, S_8$. See Figure 7.8.

Let m be a positive integer. We define the two-dimensional region

$$b_k(m) = [-m, m]^{k-1} \times \{0\} \times [-m, m]^{3-k} \quad \text{for } k = 1, 2, 3,$$

and we designate as *squares* all translates $x + b_k(m)$ for $x \in \mathbb{Z}^3$ and $k = 1, 2, 3$. A square $x + b_k(m)$ is called a *seed* if all edges joining pairs of vertices in the square are open. To each $x \in S \cup T$ we associate a square $b(x) = b(x, m, L, H)$ having x at its centre, and we do this as follows. To any $x \in T$ we associate the square $b(x) = x + b_3(m)$. If $x \in S \setminus T$, we find some i such that $x \in S_i$, and

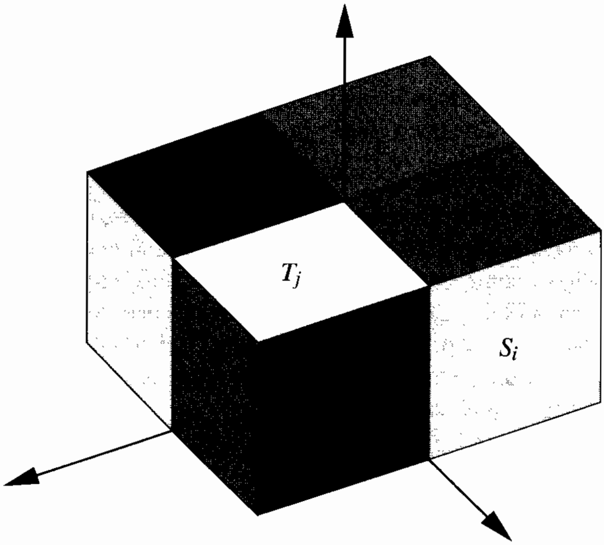


Figure 7.8. The sides and top of the brick $B(L, H)$ may be divided into subfacets S_i, T_j.

we define $k(x)$ such that $b(x) = x + b_{k(x)}(m)$ is parallel to S_i; for x belonging to more than one S_i, we pick one of these according to some predetermined rule. Finally, we let $b(0) = b_3(m)$, a square having the origin at its centre.

Suppose that $L \geq m$ and $H \geq 2m$. We call the brick $B(L, H)$ *good* if:

(a) there exist $x_i \in S_i$ for $1 \leq i \leq 8$, and $y_j \in T_j$ for $1 \leq j \leq 4$, such that every $b(x_i)$ and every $b(y_j)$ is a seed;

(b) for every i, there exists an open path of $B(L, H)$ joining $b(x_i)$ to $b(0)$ using no edges in the underside U;

(c) for every j, there exists an open path of $B(L, H)$ joining $b(y_j)$ to $b(0)$ using no edges in the underside U.

Note that this definition does not require that $b(0)$ be a seed.

The following lemma is fundamental to the block construction which follows, and is closely related to Lemma (7.9). It is valid for general p but we shall require it for $p = p_c$ only, for which case the proof is slightly shorter.

(7.36) Lemma. *Suppose that* $\theta_{\mathbb{H}}(p_c) > 0$. *If* $\eta > 0$, *there exist integers* m, L, H, *satisfying* $m \geq 1$, $L \geq m$, *and* $H \geq 2m$, *such that*

$$P_{p_c}\big(B(L, H) \text{ is good}\big) > 1 - \eta.$$

Proof. For $V \subseteq \mathbb{Z}^3$ and $x, y \in V$, we write '$x \leftrightarrow y$ in V^*' if there exists an open path in V joining x and y and using no edge joining two vertices in the boundary of V. Similarly, we write '$x \leftrightarrow \infty$ in V^*' if such an infinite open path exists having endvertex x.

Let ε satisfy $0 < \varepsilon < \frac{1}{2}$; later we shall choose ε in terms of η. Assume that $\theta_{\mathbb{H}}(p_c) > 0$, and let L denote the plane $\mathbb{Z}^2 \times \{0\}$. The event $J = \{L \leftrightarrow \infty \text{ in } \mathbb{H}^*\}$ is invariant under shifts of the form $x \mapsto x + (i, j, 0)$ for $i, j \in \mathbb{Z}$; therefore J has probability either 0 or 1, by the zero–one law for the invariant σ-field of a family of independent random variables indexed by $\mathbb{Z}^2$. Now,

$$\begin{aligned} P_{p_c}(J) &\geq P_{p_c}(0 \leftrightarrow \infty \text{ in } \mathbb{H}^*) \\ &\geq p_c P_{p_c}\big((0,0,1) \leftrightarrow \infty \text{ in } \mathbb{H} + (0,0,1)\big) \\ &= p_c \theta_{\mathbb{H}}(p_c) > 0, \end{aligned}$$

whence $P_{p_c}(J) = 1$. Therefore,

$$P_{p_c}\big(b(0) \leftrightarrow \infty \text{ in } \mathbb{H}^*\big) \to 1 \quad \text{as } m \to \infty,$$

and we choose $m = m(\varepsilon)$ such that

$$P_{p_c}\big(b(0) \leftrightarrow \infty \text{ in } \mathbb{H}^*\big) > 1 - \varepsilon^2. \tag{7.37}$$

We recall from the discussion following Theorem (3.16) that $p_c(S_h) > p_c$, where S_h is the slab $\mathbb{Z}^2 \times [0, h]$. It follows that

$$P_{p_c}\big(b(0) \leftrightarrow \infty \text{ in } S_h\big) = 0 \quad \text{for all } h \geq 0. \tag{7.38}$$

Let M be a positive integer, and set $N_1 = \lceil M(6m+1)^{-3} \rceil$. The number N_1 has been chosen in such a way that, if we are provided with M vertices $x_i \in S$ and M vertices $y_j \in T$ (these $2M$ vertices being distinct), then these include at least N_1 vertices $x \in S$ and N_1 vertices $y \in T$ such that set of all squares $b(x)$, $b(y)$ are disjoint. We choose $M = M(\varepsilon)$ sufficiently large that

$$\big\{1 - P_{p_c}(b(0) \text{ is a seed})\big\}^{N_1} < \varepsilon \tag{7.39}$$

and

$$P(Z \leq M) < \varepsilon \tag{7.40}$$

where Z has the binomial distribution with parameters $N_2 = \lceil 2M/p_c \rceil$ and p_c.

Let

$$U(h) = \big|\big\{x : x_3 = h,\ b(0) \leftrightarrow x \text{ in } S_h^*\big\}\big|.$$

As in (7.15)–(7.16) with $B(m)$ replaced by $b(0)$,

$$P_{p_c}\big(U(h) < 2N_2\big) \leq a_h + \varepsilon^2$$

by (7.37), for some a_h satisfying $a_h \to 0$ as $h \to \infty$. (We have used (7.38) surreptiously here.) We pick H_0 large enough that $a_h \le \varepsilon^2$ for $h \ge H_0$, and deduce that

$$P_{p_c}\big(U(h) \ge 2N_2\big) > 1 - \varepsilon \quad \text{if } h \ge H_0. \tag{7.41}$$

(Recall that ε was assumed smaller than $\frac{1}{2}$.)

Next, let $h \ge H_0$, and define

$$U(l,h) = \big|\{x \in T(l,h) : b(0) \leftrightarrow x \text{ in } B(l,h)^*\}\big|.$$

Since $U(l,h) \to U(h)$ as $l \to \infty$, there exists $L_0 = L_0(h)$ such that

$$P_{p_c}\big(U(l,h) \ge N_2\big) > 1 - 2\varepsilon \quad \text{if } l \ge L_0. \tag{7.42}$$

We now 'fine tune' the height h in such a way that $U(l,h)$ is 'only just large enough'; this will enable the conclusion that there exist (with large probability) many points in both $S(l,h)$ and $T(l,h)$ which are joined to $b(0)$ in $B(l,h)^*$. Suppose that $l \ge L_0(H_0)$. The sequence $a(h) = P_{p_c}(U(l,h) \ge N_2)$ satisfies $a(H_0) > 1 - 2\varepsilon$ (by (7.42)), and furthermore

$$\begin{aligned} a(h) &\le P_{p_c}\big(b(0) \leftrightarrow T(l,h) \text{ in } B(l,h)\big) \\ &\to P_{p_c}\big(b(0) \leftrightarrow \infty \text{ in } B(l,\infty)\big) \quad \text{as } h \to \infty. \end{aligned}$$

The last probability equals 0 since $B(l,\infty)$ is topologically one-dimensional. It follows that there exists an integer $H_1 = H_1(l)$ $(> H_0)$ such that

$$P_{p_c}\big(U(l, H_1 - 1) \ge N_2\big) > 1 - 2\varepsilon \ge P_{p_c}\big(U(l, H_1) \ge N_2\big). \tag{7.43}$$

Since $U(l,h) \uparrow U(h)$ as $l \to \infty$, we have that $H_1(l)$ is non-decreasing in l. Let

$$H_1(\infty) = \lim_{l \to \infty} H_1(l).$$

If $H_1(\infty) < \infty$, it follows by (7.43) and the fact that $U(l, H_1(l)) \to U(H_1(\infty))$ that

$$P_{p_c}\big(U(H_1(\infty)) \ge N_2\big) \le 1 - 2\varepsilon,$$

in contradiction of (7.41). Therefore,

$$H_1(l) \to \infty \quad \text{as } l \to \infty. \tag{7.44}$$

We now introduce the random variable

$$V(l,h) = \big|\{x \in S(l,h) : b(0) \leftrightarrow x \text{ in } B(l,h)^*\}\big|.$$

Our next target is to show that we can choose the pair l, h in such a way that, with large probability, both $U(l,h)$ and $V(l,h)$ are large. First, if $l \geq L_0(H_0)$,

(7.45)
$$\begin{aligned} P_{p_c}\big(U(l, H_1(l)) \geq M\big) &\\ \geq P_{p_c}\big(U(l, H_1(l)-1) \geq N_2\big) &\\ \times P_{p_c}\big(U(l, H_1(l)) \geq M \,\big|\, U(l, H_1(l)-1) \geq N_2\big) &\\ > (1-2\varepsilon) P_{p_c}\big(U(l, H_1(l)) \geq M \,\big|\, U(l, H_1(l)-1) \geq N_2\big), & \end{aligned}$$

by (7.43). Now, given N_2 points z in $T(l, H_1(l)-1)$, there is by (7.40) at least probability $1-\varepsilon$ that M or more of the edges $\langle z, z+(0,0,1)\rangle$ are open (note that this last event is conditionally independent of the states of all edges in $B(l, H_1(l)-1)$). Therefore, the conditional probability in (7.45) is at least $1-\varepsilon$, whence

(7.46) $$P_{p_c}\big(U(l, H_1(l)) \geq M\big) > 1 - 3\varepsilon \quad \text{for } l \geq L_0(H_0).$$

We next prove a similar property for $V(l, H_1(l))$. It is a consequence of (7.44) that the bricks $B(l, H_1(l))$ satisfy $B(l, H_1(l)) \uparrow \mathbb{H}$ as $l \to \infty$. As in (7.15)–(7.16) with $U(n)$ replaced by $U(l, H_1(l)) + V(l, H_1(l))$ and $B(m)$ replaced by $b(0)$,

$$P_{p_c}\big(U(l, H_1(l)) + V(l, H_1(l)) < M + N_2\big) \leq a_l + \varepsilon^2$$

for some a_l satisfying $a_l \to 0$ as $l \to \infty$. We may therefore find $L_1 = L_1(H_0) > L_0(H_0)$ such that

$$P_{p_c}\big(U(l, H_1(l)) + V(l, H_1(l)) < M + N_2\big) \leq 2\varepsilon^2 \quad \text{for } l \geq L_1.$$

It follows by the FKG inequality that

$$\begin{aligned} P_{p_c}\big(U(l, H_1(l)) < N_2\big) P_{p_c}\big(V(l, H_1(l)) < M\big) &\\ \leq P_{p_c}\big(U(l, H_1(l)) + V(l, H_1(l)) < M + N_2\big) &\\ \leq 2\varepsilon^2 \quad \text{if } l \geq L_1, & \end{aligned}$$

whence, by (7.43),

(7.47) $$P_{p_c}\big(V(l, H_1(l)) \geq M\big) > 1 - \varepsilon \quad \text{if } l \geq L_1.$$

Set $L = L_1$ and $H = H_1(L_1)$, so that $U = U(L, H)$ and $V = V(L, H)$ satisfy

(7.48) $$P_{p_c}(U \geq M) > 1 - 3\varepsilon, \quad P_{p_c}(V \geq M) > 1 - \varepsilon$$

by (7.46)–(7.47). If $U \geq M$ then, by the discussion before (7.39), there exists some subset $\mathcal{N}$ of $T(L, H)$ such that:

(i) $|\mathcal{N}| \geq N_1$;
(ii) for all $x \in \mathcal{N}$; $b(0) \leftrightarrow x$ in $B(L, H)^*$;
(iii) the squares $b(x)$, $x \in \mathcal{N}$, are disjoint;
(iv) the set $\mathcal{N}$ is measurable on the σ-field generated by the set of states of edges having at least one endvertex in the interior of $B(L, H)$.

Since the squares $b(x)$, $x \in \mathcal{N}$, do not intersect the interior of $B(L, H)$, we deduce by (7.39) that

$$P_{p_c}\big(X = 0 \,\big|\, U \geq M\big) < \varepsilon,$$

where

$$X = \Big|\big\{x \in T(L, H) : b(0) \leftrightarrow x \text{ in } B(L, H)^*,\ b(x) \text{ is a seed}\big\}\Big|.$$

In combination with (7.48), this yields

$$P_{p_c}(X = 0) \leq P_{p_c}(U < M) + \varepsilon P_{p_c}(U \geq M) < 4\varepsilon. \tag{7.49}$$

A similar argument applied to

$$Y = \Big|\big\{x \in S(L, H) : b(0) \leftrightarrow x \text{ in } B(L, H)^*,\ b(x) \text{ is a seed}\big\}\Big|$$

yields

$$P_{p_c}(Y = 0) < 2\varepsilon. \tag{7.50}$$

Inequalities (7.49)–(7.50) are not quite good enough for our purpose, since the property of being good requires seeds in every 'subfacet' S_i of S and T_j of T. This we shall obtain by further applications of the FKG inequality. Let

$$X_j = \Big|\big\{x \in T_j : b(0) \leftrightarrow x \text{ in } B(L, H)^*,\ b(x) \text{ is a seed}\big\}\Big|, \quad j = 1, 2, 3, 4,$$
$$Y_i = \Big|\big\{x \in S_i : b(0) \leftrightarrow x \text{ in } B(L, H)^*,\ b(x) \text{ is a seed}\big\}\Big|, \quad i = 1, 2, \dots, 8.$$

By symmetry, the X_j (respectively Y_i) have the same distribution, and in addition

$$X \leq \sum_j X_j, \quad Y \leq \sum_i Y_i.$$

Therefore,

$$\begin{aligned} P_{p_c}(X_1 = 0)^4 &= \prod_j P_{p_c}(X_j = 0) \\ &\leq P_{p_c}(X_j = 0 \text{ for all } j) \\ &\leq P_{p_c}(X = 0) < 4\varepsilon \end{aligned}$$

by the FKG inequality and (7.49), whence

$$P_{p_c}(X_j = 0) < (4\varepsilon)^{1/4} \quad \text{for } j = 1, 2, 3, 4.$$

Similarly,

$$P_{p_c}(Y_i = 0) < (2\varepsilon)^{1/8} \quad \text{for } i = 1, 2, \ldots, 8. \tag{7.51}$$

Let $\eta > 0$, and choose ε such that

$$\eta = 4(4\varepsilon)^{1/4} + 8(2\varepsilon)^{1/8}.$$

It follows from the above that

$$\begin{aligned} P_{p_c}\big(B(L, H) \text{ is not good}\big) \\ &\le P_{p_c}\big(\text{either } X_j = 0 \text{ for some } j, \text{ or } Y_i = 0 \text{ for some } i\big) \\ &< 4(4\varepsilon)^{1/4} + 8(2\varepsilon)^{1/8} = \eta, \end{aligned}$$

as required. □

Our next lemma will be proved by a construction similar to that of the last section.

(7.52) Lemma. *There exists a strictly positive number ν such that the following holds. Let $0 < p < 1$. Suppose that m, L, H are positive integers satisfying $m \ge 1$, $L \ge m$, $H \ge 2m$, and that*

$$P_p\big(B(L, H) \text{ is good}\big) > 1 - \nu. \tag{7.53}$$

Then $\theta_{\mathbb{H}}(p) > 0$.

Theorem (7.35) is a consequence of Lemmas (7.36) and (7.52), as follows. Suppose that $\theta_{\mathbb{H}}(p_c) > 0$, and let ν be given as in Theorem (7.52). By Lemma (7.36), there exist integers m, L, H satisfying $m \ge 1$, $L \ge m$, $H \ge 2m$, such that (7.53) holds with $p = p_c$. The event $\{B(L, H) \text{ is good}\}$ depends on the states of a finite set of edges only, whence $P_p(B(L, H) \text{ is good})$ is a continuous function of p. Therefore, there exists p' satisfying $p' < p_c$ such that

$$P_{p'}\big(B(L, H) \text{ is good}\big) > 1 - \nu.$$

We have by Lemma (7.52) that $\theta_{\mathbb{H}}(p') > 0$, and it follows by contradiction that $\theta_{\mathbb{H}}(p_c) = 0$.

Proof of Lemma (7.52). Let $m \ge 1$, $L \ge m$, $H \ge 2m$, $0 < p < 1$, and set

$$\pi = P_p\big(B(L, H) \text{ is good}\big).$$

We shall use rotated translates of $B(L, H)$ in a attempt to build an infinite open path within a half-space of $\mathbb{Z}^3$. Such a construction will succeed with a strictly positive probability so long as π is sufficiently close to 1. We do not present all the details of the construction, referring the insistent reader to Barsky, Grimmett, and Newman (1991b). Instead, we present the salient features only of a version of the required argument.

We divide $\mathbb{Z}^3$ into infinite tubes having width $2N$ for an appropriate value of N. An unfortunate complication is that we shall require different constructions depending on whether or not $H \geq L$. Let

$$N = \begin{cases} 5H + 7L & \text{if } H \geq L, \\ 11H + 11L & \text{if } H < L, \end{cases} \tag{7.54}$$

and define the tubes τ_a, indexed by the set $\mathbb{Z}_+ \times \mathbb{Z}$, by

$$\begin{aligned} \tau_0 &= \mathbb{Z} \times [-N, N]^2, \\ \tau_a &= 2N(0, a_2, a_3) + \tau_0 \quad \text{for } a = (a_2, a_3) \in \mathbb{Z}_+ \times \mathbb{Z}. \end{aligned}$$

A pair τ_a, τ_b of tubes is deemed *adjacent* if $a \sim b$.

Four subsets of τ_0 are designated *target zones*, namely

$$\begin{aligned} Z_0^1 &= \mathbb{Z} \times [-\alpha, \alpha] \times [\beta, N], & Z_0^2 &= \mathbb{Z} \times [\beta, N] \times [-\alpha, \alpha], \\ Z_0^3 &= \mathbb{Z} \times [-\alpha, \alpha] \times [-N, -\beta], & Z_0^4 &= \mathbb{Z} \times [-N, -\beta] \times [-\alpha, \alpha], \end{aligned}$$

where

$$\begin{aligned} \alpha &= 4H + 5L, \ \beta = N - H + 1, && \text{if } H \geq L, \\ \alpha &= 9H + 9L, \ \beta = N - 2H - L + 1, && \text{if } H < L. \end{aligned}$$

Similarly, for $a = (a_2, a_3) \in \mathbb{Z}_+ \times \mathbb{Z}$, the regions $Z_a^i = Z_0^i + 2N(0, a_2, a_3)$, $1 \leq i \leq 4$, are termed the target zones of the tube τ_a. Note that the target zones of τ_a do not intersect one another, and that they abut the four faces of τ_a that intersect the four neighbouring tubes τ_b with $b \in a + \{(1, 0), (0, 1), (-1, 0), (0, -1)\}$. See Figure 7.9.

We shall examine tubes one by one. On examining the tube τ_a, we shall declare the vertex a ($\in \mathbb{Z}_+ \times \mathbb{Z}$) to be either 'occupied' or 'unoccupied' depending on the existence (or not) of certain open paths. This will be done in such a way that:

(a) if π is sufficiently large, there is conditional probability strictly exceeding $p_c^{\text{site}}(\mathbb{Z}_+ \times \mathbb{Z})$ that an examined vertex is occupied;

(b) the existence of an infinite occupied cluster in $\mathbb{Z}_+ \times \mathbb{Z}$ implies the existence of an infinite open path within the corresponding region $\mathbb{Z} \times [-N, \infty) \times \mathbb{Z}$ of $\mathbb{Z}^3$.

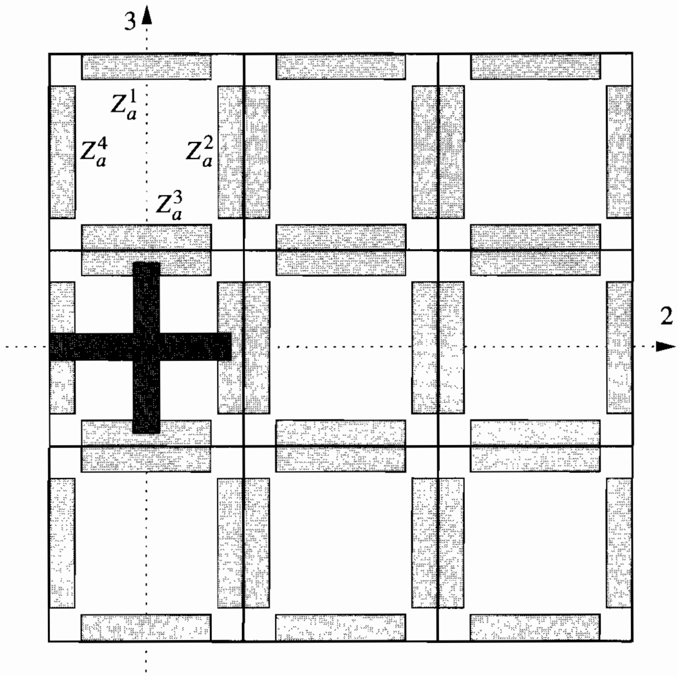


Figure 7.9. A map of $\mathbb{L}^3$ showing cross-sections of the tubes τ_a and their target zones. The axes are labelled by the appropriate coordinate number. The darker zone is the region E.

Since the latter region is a half-space of $\mathbb{Z}^3$, this will imply as required that $\theta_{\mathbb{H}}(p) > 0$ when π is large. [Note that $p_c^{\text{site}}(\mathbb{Z}_+ \times \mathbb{Z}) < 1$. This elementary fact may be proved in a variety of ways, of which the most obvious is an adaptation of the method used to prove Theorem (1.10).]

The algorithm to be followed is that utilized in the proofs of Theorems (1.33) and (7.2). First we examine the tube τ_0, and we decide whether or not 0 is occupied, according to a certain rule to be specified. If 0 is unoccupied, we cease the construction. If 0 is occupied, we find its earliest neighbour a ($\in \mathbb{Z}_+ \times \mathbb{Z}$) (in some predetermined ordering of vertices) and we consider τ_a. Continuing likewise, we build up an occupied cluster of vertices in $\mathbb{Z}_+ \times \mathbb{Z}$, and a partial boundary of unoccupied vertices. Next, we explain what is meant by the term 'occupied', beginning with the slightly special case of the first tube τ_0.

Let E be the set of all edges of $\mathbb{L}^3$ joining pairs of vertices in the region

$$\Big([-m, m]^2 \times [-N+m, N-m]\Big) \cup \Big([-m, m] \times [-N, N-m] \times [-m, m]\Big);$$

see Figure 7.9. We declare 0 to be occupied if every edge in E is open, noting that

$$(7.55) \qquad P_p(0 \text{ is occupied}) > 0.$$

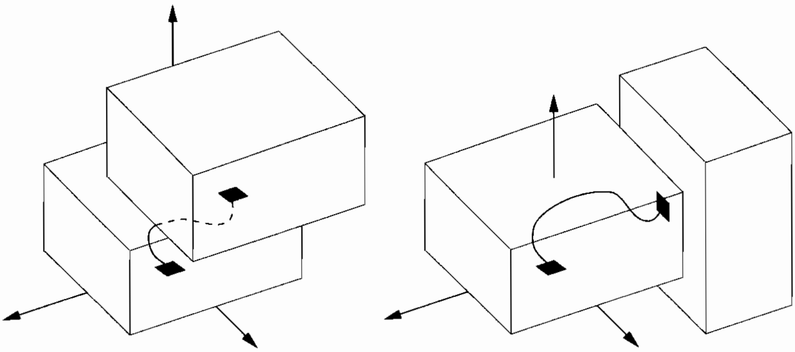


Figure 7.10. Illustrations of a top stacking and a side stacking.

If 0 is not occupied, we call it 'unoccupied'. When 0 is occupied, the squares

$$[-m, m]^2 \times \{N - m\},$$
$$[-m, m]^2 \times \{-N + m\},$$
$$[-m, m] \times \{N - m\} \times [-m, m],$$

are seeds to which the vertex $(0, -N, 0)$ is connected by open paths of E. These seeds may be used to initiate further construction.

Assume that 0 is occupied, and find the earliest neighbour of 0 in $\mathbb{Z}_+ \times \mathbb{Z}$; for the sake of definiteness, we assume this to be the vertex $a = (0, 1)$. Our target is to extend the open region E into τ_a by a careful stacking of rotated translates of the brick $B(L, H)$. The following notes will aid an understanding of how this stacking is to be achieved.

(A) *Top stacking.* If $B(L, H)$ is good, then $b(0)$ is connected by open paths to seeds in each subfacet S_i, T_j. Let j be given, and suppose that $b(y_j)$ is such a seed with $y_j \in T_j$. If, further, the translate $y_j + B(L, H)$ is good, then there exist concatenated open paths joining $b(0)$ to seeds on all subfacets of $y_j + B(L, H)$. Given that $B(L, H)$ is good, we may choose such a y_j according to some fixed ordering of all available such vertices; there is conditional probability π that $y_j + B(L, H)$ is good, since this event is defined in terms of edges lying outside $B(L, H)$. Therefore, there is probability π^2 that this 'top stacking' of two bricks succeeds. Similarly, a 'top stacking' of n bricks succeeds with probability π^n. See the illustration on the left of Figure 7.10.

(B) *Side stacking.* Top stacking, if repeated without limit, will almost surely fail sooner or later since $\pi < 1$; therefore, we shall need to stack on the sides of good bricks as well as on their tops. We illustrate this by an example. Suppose that the brick $y + B(L, H)$ has been found to be good. Let S_i be a subfacet of the sides of $B(L, H)$; for the sake of definiteness, we assume that $S_i \subseteq \{x \in \mathbb{Z}^3 : x_2 = L\}$. We may find $x_i \in S_i$ such that $y + b(0)$ is

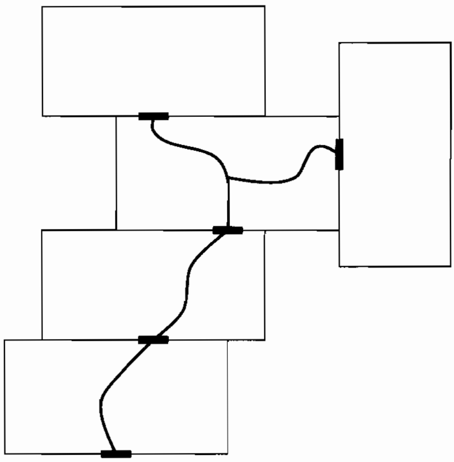


Figure 7.11. By a careful choice of subfacets, one may stack on both top and sides of an occupied brick. The short black lines on the undersides of the bricks represent seeds.

joined to $y + x_i$ in $y + B(L, H)^*$, and that $y + b(x_i)$ is a seed. We then consider the brick

$$B' = y + x_i + [-L, L] \times [0, H] \times [-L, L]$$

obtained by a rotation and translation of $B(L, H)$. If the intersection of B' with the region of space examined so far is limited to a subset of its underside $y + x_i + [-L, L] \times \{0\} \times [-L, L]$, then B' is good with (conditional) probability π. If indeed B' is good, then the construction may continue as before by stacking further bricks on the top of B'. See the illustration on the right of Figure 7.10.

(C) *Branching.* In a similar way, one may stack bricks in a variety of different directions on top of a given good brick. In order that we may control the (conditional) probabilities of success, such new bricks should be stacked one by one, and at every stage the intersection of a new brick with the region considered so far must be limited to a subset of its underside. We shall never need to stack more than one brick on the top (respectively sides) of a given good brick, but we shall indeed require the ability to stack bricks on the top as well as on the sides. At this point we shall use our freedom to choose seeds within particular subfacets of the top and sides; by choosing such subfacets with care, one may ensure that these two bricks do not intersect one another. See Figure 7.11.

(D) *Steering.* If we were to stack bricks one by one on top of one another, the pile would grow in one direction but would generally deviate substantially in the other two coordinate directions. By judicious choices of the particular subfacets of the tops of the bricks, we may control such deviations. For example, suppose that the pile is growing in the direction of increasing first coordinate, but that it deviates towards the direction of increasing second coordinate. We may rectify this by choosing a subfacet of the top containing points with negative second coordinates.

We return now to the main argument. Suppose that 0 is occupied, and consider the vertex $a = (0, 1)$ and associated tube τ_a. Since 0 is occupied, the square $[-m, m]^2 \times \{N - m\}$ is a seed. Consider the brick $B' = (0, 0, N - m) + B(L, H)$ having this seed at the centre of its underside. If B' is good, we stack a brick B'' on its top according to (A) above, and we ask whether B'' is good. If so, then we iterate the construction until the first time that a new brick intersects the target zone Z_a^1. In so doing, we employ the steering mechanism described in (D) above in order that the pile of blocks remain close to the 'mid-line' $\{x \in \mathbb{Z}^3 : x_1 = x_2 = 0\}$. We do not give the geometrical details of this, but we note that the bricks may be chosen in such a way that the pile deviates only a little from this mid-line. Furthermore, we may position the bricks in this growing pile in such a way as to permit a branching 'westwards' and 'eastwards' (such compass directions refer to the obvious map of $\mathbb{Z}_+ \times \mathbb{Z}$, as in Figure 7.9). Having branched as in (C) above, the westerly and easterly branches are 'steered' into the target zones Z_a^4 and Z_a^2 respectively. If all three target zones are reached in this construction, we declare the vertex a to be occupied; otherwise, we declare a unoccupied.

If 0 and a are both occupied, the vertex $(0, -N, 0)$ is connected by open paths contained within the good bricks to seeds in each of the target zones Z_0^i, Z_a^i, for $i = 1, 2, 3, 4$. Some of these seeds will be used to initiate further stages of the construction.

There are several details to be checked in verifying this construction. In order to control the success probability of a tube in terms of the quantity π, we need to employ a construction with the following property: there exists an absolute upper bound R on the number of new bricks required for a vertex of $\mathbb{Z}_+ \times \mathbb{Z}$ to be occupied. It is for this reason that the actual construction depends on whether $H < L$ or $H \geq L$. The exact details of the constructions necessary for these two cases have appeared in Barsky, Grimmett, and Newman (1991b), and will not be repeated here. They are illustrated in Figure 7.12.

With N given in (7.54), it turns out that one may take $R = 125$. Therefore,

$$(7.56)\quad P_p\big(a \text{ is occupied} \,\big|\, 0 \text{ is occupied}\big) \geq \Big[P_p\big(B(L, H) \text{ is good}\big)\Big]^{125} = \pi^{125}.$$

The general step is very similar. Suppose we wish to determine whether some vertex c ($\in \mathbb{Z}_+ \times \mathbb{Z}$) is occupied. We find a neighbouring b such that b has already been determined to be occupied. The target zone of τ_b which abuts τ_c has been reached by a successful construction. We seek to extend this construction in the manner described above, thereby reaching the three 'new' target zones of τ_c. If we succeed in this, then we declare c to be occupied; otherwise, we declare c

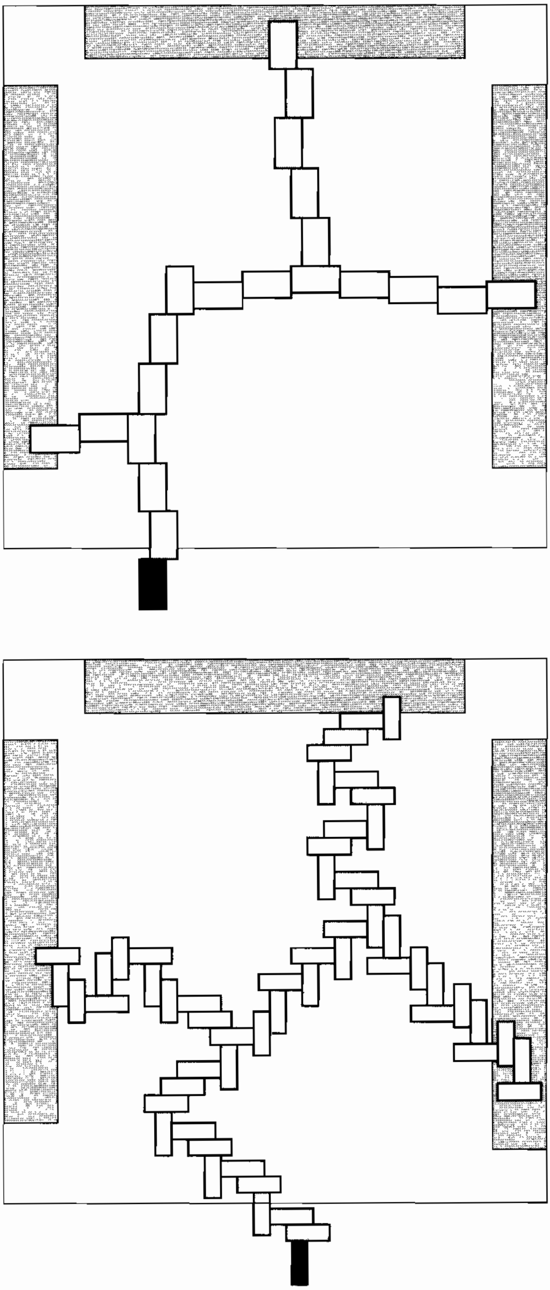


Figure 7.12. Illustrations of the block constructions for the proof of Lemma (7.52). The towers of bricks contain open paths joining the black 'inlet' brick to the three shaded target zones. We have no control over the aspect ratio of the basic brick, and consequently two cases with somewhat different geometries need to be considered. This figure may be compared with Figures 7.5 and 7.6.

unoccupied. As in (7.56), each new step of this process results in an occupied vertex with conditional probability at least π^{125}. Furthermore, every time that a vertex is found to be occupied, we have successfully constructed certain open paths joining the vertex $(0, -N, 0)$ to the target zones of the corresponding tube.

If $\pi^{125} > p_c^{\text{site}}(\mathbb{Z}_+ \times \mathbb{Z})$, there exists a strictly positive probability that 0 lies in an infinite cluster of occupied vertices of $\mathbb{Z}_+ \times \mathbb{Z}$; this holds just as in Lemma (7.24). The existence of such an infinite cluster implies that the vertex $(0, -N, 0)$ lies in an infinite open path of the half-space $\mathbb{Z} \times [-N, \infty) \times \mathbb{Z}$, whence $\theta_{\mathbb{H}}(p) > 0$. In summary, if (7.53) holds with ν given by

$$(1 - \nu)^{125} = p_c^{\text{site}}(\mathbb{Z}_+ \times \mathbb{Z}),$$

then $\theta_{\mathbb{H}}(p) > 0$. The lemma has been proved. □

7.4 Static Renormalization

The block arguments described in this chapter so far have been 'dynamic' in the following sense. At each stage of an iterative process, we have encountered a new region R of space, and we have constructed an associated event E_R. We have then arranged that E_R have conditional probability close to 1. The events E_R have been defined in terms of the states of edges encountered earlier, and it is in this sense that the process is 'dynamic'. The inter-dependence of the events E_R can be a hindrance in applications, and the target of this section is to describe a type of 'static' renormalization which can be an easier tool in certain settings. There are several ways of doing this, of which we shall employ the following.

Let n be a positive integer, and let $B(n) = [-n, n]^d$ as usual. The open edges of $B(n)$ form open clusters in $B(n)$, and we write $M(n)$ for the largest such cluster (if there exist two or more largest clusters, we pick one according to some predetermined rule). We say that $M(n)$ 'crosses $B(n)$ in the ith direction' if $M(n)$ contains an open path π having endvertices x, y satisfying $x_i = -n$, $y_i = n$. We call $M(n)$ a *crossing cluster* of $B(n)$ if $M(n)$ crosses $B(n)$ in the ith direction for $i = 1, 2, \ldots, d$. Crossing clusters of other boxes in $\mathbb{L}^d$ are defined similarly.

Here is some further notation. For $A \subseteq \mathbb{Z}^d$, let

$$L_i = \inf\{x_i : x \in A\}, \qquad R_i = \sup\{x_i : i \in A\},$$

and define the *diameter* $\operatorname{diam}(A)$ by

$$\operatorname{diam}(A) = \max\{R_i - L_i : 1 \leq i \leq d\}.$$

We note that every crossing cluster of $B(n)$ has diameter $2n$.

Let $0 < \varepsilon < 1$, $0 \le p \le 1$, and $\omega \in \Omega$. We declare the region $B(n)$ to be *ε-good* if:

(a) $M(n)$ is a crossing cluster of $B(n)$;

(b) $M(n)$ is the unique open cluster C of $B(n)$ satisfying $\text{diam}(C) \ge n$; and

(c) $|M(n)| \ge (1-\varepsilon)\theta(p)|B(n)|$.

We make three remarks concerning this definition. First, the event that $B(n)$ is ε-good is defined in terms of the states of edges in $B(n)$. Secondly, condition (b) may be replaced by:

(b′) $M(n)$ is the unique open cluster C of $B(n)$ satisfying $\text{diam}(C) \ge \nu(p)\log n$,

where $\nu(p)$ is suitably large. The proofs which follow are essentially the same with (b′) in place of (b).

Finally, requirement (c) will not be used in the rest of this book, but we include it here since it can be a desirable property for certain applications. Only in (c) is the value of ε relevant. One may replace (c) by a variety of requirements concerning the open clusters of $B(n)$, so long as the corresponding event has probability tending to 1 as $n \to \infty$. Further discussion of this point may be found in the notes.

We now extend this definition to translates of $B(n)$. For $x \in \mathbb{Z}^d$, let

$$B_x = B_x(n) = nx + B(n), \tag{7.57}$$

and let $M_x(n)$ denote the largest open cluster of $B_x(n)$. We call $B_x(n)$ *ε-good* if (a), (b), (c) hold with $M(n)$ and $B(n)$ replaced by $M_x(n)$ and $B_x(n)$. We now designate the vertex x ($\in \mathbb{Z}^d$) *ε-occupied* if $B_x(n)$ is ε-good; we designate it *ε-unoccupied* otherwise. Such designations give rise to a family $X_\varepsilon = \{X_{\varepsilon,x} : x \in \mathbb{Z}^d\}$ of $\{0,1\}$-valued random variables, given by

$$X_{\varepsilon,x} = X_{\varepsilon,x}(n) = \begin{cases} 1 & \text{if } x \text{ is } \varepsilon\text{-occupied,} \\ 0 & \text{otherwise.} \end{cases} \tag{7.58}$$

About the 'block variables' $X_{\varepsilon,x}$ we note two facts, the first of which is geometrical and the second probabilistic.

Consider two neighbours of $\mathbb{Z}^d$, say

$$x = (x_1, x_2, \dots, x_d) \text{ and } y = (x_1, x_2, \dots, x_{i-1}, x_i + 1, x_{i+1}, \dots, x_d),$$

and suppose that $X_{\varepsilon,x} = X_{\varepsilon,y} = 1$. We have that $M_y(n)$ crosses $B_y(n)$ in the ith direction, and therefore contains an open path having diameter at least n lying within the intersection

$$B_x(n) \cap B_y(n) = x + \left\{[-n,n]^{i-1} \times [0,n] \times [-n,n]^{d-i}\right\}. \tag{7.59}$$

Since $X_{\varepsilon,x} = 1$, this path lies in the largest open cluster $M_x(n)$ of $B_x(n)$, and we conclude that $M_x(n)$ and $M_y(n)$ have non-empty intersection. (See Figure 7.13.)

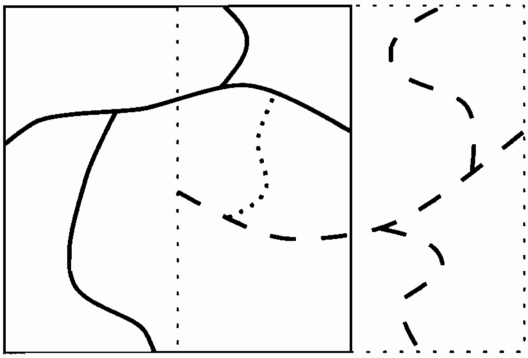


Figure 7.13. If x and y are ε-occupied neighbours, the corresponding boxes contain crossing clusters, and these clusters are connected to one another in the union of the boxes.

More generally, if $s(1), s(2), \dots, s(k)$ is a path in $\mathbb{Z}^d$ with $X_{\varepsilon,s(i)} = 1$ for all i, the corresponding largest clusters $M_{s(i)}(n)$, $1 \le i \le k$, belong to the same open cluster of the corresponding region $\bigcup_i B_{s(i)}(n)$ of the original lattice. In this way, clusters of ε-occupied vertices correspond to connected open subgraphs of the original lattice. This geometrical property is central to applications of the ensuing block argument.

Having defined the events in question, we turn to probabilities. We note first that the family $X_\varepsilon = \{X_{\varepsilon,x} : x \in \mathbb{Z}^d\}$ is stationary under lattice shifts. It follows in particular that

(7.60) $$P_p(X_{\varepsilon,x} = 1) = P_p(X_{\varepsilon,0} = 1) \quad \text{for all } x \in \mathbb{Z}^d.$$

A family $Y = \{Y_x : x \in \mathbb{Z}^d\}$ is called *k-dependent* if any two sub-families $\{Y_x : x \in A\}$ and $\{Y_x : x \in A'\}$ are independent whenever $\delta(x, y) > k$ for all $x \in A$, $y \in A'$. [Recall that $\delta(\cdot, \cdot)$ denotes graph-theoretic distance.] The random variable $X_{\varepsilon,x}$ is defined in terms of the states of edges inside $B_x(n)$, and $B_x(n) \cap B_y(n) = \varnothing$ whenever $\delta(x, y) > 3d$. It follows that $X_\varepsilon = \{X_{\varepsilon,x} : x \in \mathbb{Z}^d\}$ is a $3d$-dependent family.

We shall require two theorems in order to study the $X_{\varepsilon,x}$. The first says that, when $p > p_c$, the density of the $X_{\varepsilon,x}$ may be made as large as required, by means of a suitable choice of n. The second implies that, for large n, the family X_ε dominates (stochastically) a family of independent random variables having large density. The latter family may be studied as a supercritical site percolation model, and this fact will be used, via the intermediate 'block variables' $X_{\varepsilon,x}$, in order to understand the geometry of the original percolation process.

(7.61) Theorem. *Let $d \ge 2$, $0 < \varepsilon < 1$, and $p > p_c$. It is the case that*

(7.62) $$P_p\big(X_{\varepsilon,0}(n) = 1\big) \to 1 \quad \textit{as } n \to \infty.$$

We turn now to the required comparison inequality, and we begin with a more systematic account of stochastic domination than has been presented so far. Let S be a countable set, and let $Y = \{Y_x : x \in S\}$, $Z = \{Z_x : x \in S\}$ be families of random variables taking values in the set $\{0, 1\}$ and indexed by S. We say that Y *dominates (stochastically)* Z, written $Y \geq_{\text{st}} Z$, if

$$E(f(Y)) \geq E(f(Z)) \tag{7.63}$$

for all bounded, increasing, measurable functions $f : \{0, 1\}^S \to \mathbb{R}$. [Here and later, we use E and P to denote the expectation operator and probability measure appropriate for the setting.] We shall be concerned largely with the case when Z is a family of independent random variables, and we introduce some notation to facilitate this. For $0 \leq \pi \leq 1$, we let $Z^\pi = \{Z_x^\pi : x \in S\}$ be a family of independent random variables satisfying

$$P(Z_x^\pi = 1) = 1 - P(Z_x^\pi = 0) = \pi \quad \text{for all } x \in S.$$

It is straightforward to prove the following. Let $s_1, s_2, \ldots$ be an ordering of S, and assume that

$$P\big(Y_{s_{j+1}} = 1 \,\big|\, Y_{s_i} = y_i \text{ for } 1 \leq i \leq j\big) \geq \pi \tag{7.64}$$

for all j and all vectors $y_1, y_2, \ldots, y_j$ for which

$$P(Y_{s_i} = y_i \text{ for } 1 \leq i \leq j) > 0.$$

Under this assumption, we have that $Y \geq_{\text{st}} Z^\pi$.

(7.65) Theorem. *Let $d, k \geq 1$. There exists a non-decreasing function $\pi : [0, 1] \to [0, 1]$ satisfying $\pi(\delta) \to 1$ as $\delta \to 1$ such that the following holds. If $Y = \{Y_x : x \in \mathbb{Z}^d\}$ is a k-dependent family of random variables satisfying*

$$P(Y_x = 1) \geq \delta \quad \textit{for all } x \in \mathbb{Z}^d, \tag{7.66}$$

then

$$Y \geq_{\text{st}} Z^{\pi(\delta)}. \tag{7.67}$$

That is to say, any k-dependent family having sufficiently large density dominates an independent family having high density.

We give an application to percolation of Theorems (7.61) and (7.65) in advance of their proofs. Of the several possible such applications, we choose one concerning the number of open crossings of a large box. A *left–right crossing* of the box $B(r)$ is a path of $B(r)$ whose endvertices x, y satisfy $x_1 = -r$, $y_1 = r$. Let Λ_r be the maximal number of edge-disjoint open left–right crossings of $B(r)$.

(7.68) Theorem. *Let $d \geq 3$ and $p > p_c$. There exist strictly positive quantities $\beta = \beta(p)$ and $\gamma = \gamma(p)$ such that*

(7.69) $$P_p\big(\Lambda_r \geq \beta r^{d-1}\big) \geq 1 - \exp\{-\gamma r^{d-1}\} \quad \textit{for } r \geq 1.$$

A two-dimensional version of this theorem will be given in Theorem (11.22), and indeed the proof of (7.69) is based on the arguments leading to (11.22). The moral of this example is that two-dimensional results may be used, via block arguments, to obtain corresponding results valid for all $d \geq 2$.

Proof of Theorem (7.68). The inequality corresponding to (7.69) for *bond* percolation on $\mathbb{L}^2$ is presented in Theorem (11.22), and we take this as a starting point for the current proof. The argument used to prove Theorem (11.22) is easily adapted to obtain the following rather weak result for two-dimensional site percolation. Let $d = 2$, and consider site percolation on $\mathbb{L}^2$. Let A_s be the event that there exists an open left–right crossing of the two-dimensional box $[-s, s]^2$. There exists $\alpha \in (0, 1)$ such that the following holds. There exists $\rho = \rho(p)$ satisfying $\rho(p) > 0$ when $p > \alpha$, such that

(7.70) $$P_p(A_s) \geq 1 - e^{-\rho s} \quad \text{for } s \geq 1.$$

The general idea of the present proof is to partition $B(r)$ into two-dimensional slabs, and to show via (7.70) that many of these contain left–right crossings of $B(r)$.

Let $d \geq 3$, and take p, p' satisfying $p > p' > p_c$. We may apply Theorem (7.65) with $k = 3d$ to find a value of δ (< 1) such that $\pi(\delta) > \alpha$. Let $0 < \varepsilon < 1$, and apply Theorem (7.61) to find a positive integer N such that

$$P_{p'}\big(X_{\varepsilon,x}(N) = 1\big) > \delta \quad \text{for } x \in \mathbb{Z}^d.$$

The actual value of ε is immaterial for this application.

Let us suppose that r is a positive integer of the form

(7.71) $$r = N(K + 1)$$

for some positive integer K; a minor variation of the following argument is valid for more general integers r. For each

$$k = (k_3, k_4, \ldots, k_d) \in [-K, K]^{d-2},$$

we define

(7.72) $$S(k) = \big\{x \in \mathbb{Z}^d : |x_1|, |x_2| \leq K,\ x_i = k_i \text{ for } i \geq 3\big\}.$$

We note two facts about the $S(k)$: first, they are disjoint regions of $\mathbb{Z}^d$, and secondly, each is isomorphic to the two-dimensional box $[-K, K]^2$.

Let R_k be the event that there exists an ε-occupied left–right crossing of $S(k)$, that is, an ε-occupied path whose endvertices u, v satisfy $u_1 = -K$, $v_1 = K$. Since every x is ε-occupied with probability exceeding δ, we have from Theorem (7.65) that the set of ε-occupied vertices in $S(k)$ dominates (stochastically) a site percolation process on $S(k)$ having density $\pi(\delta)$ ($> \alpha$). Since R_k is an increasing event, we have by (7.70) that

$$P_{p'}(R_k) \geq P_{\pi(\delta)}(A_K) \geq 1 - e^{-\rho K}, \tag{7.73}$$

where $\rho = \rho(\pi(\delta)) > 0$. On the event R_k, the region $NS(k) + B(N)$ of the original lattice contains a left–right crossing of $B(r)$.

Let

$$\mathcal{K} = \left\{k \in [-K, K]^{d-2} : \text{ the final component of } k \text{ is divisible by } 3\right\}.$$

The regions $\{NS(k) + B(N) : k \in \mathcal{K}\}$, are disjoint, whence the events $\{R_k : k \in \mathcal{K}\}$, are independent. Therefore,

$$\begin{aligned} P_{p'}(\Lambda_r = 0) &\leq P_{p'}\left(\bigcap_{k\in\mathcal{K}} \overline{R_k}\right) \\ &= \prod_{k\in\mathcal{K}} P_{p'}(\overline{R_k}) \leq \exp\{-\rho K|\mathcal{K}|\} \end{aligned} \tag{7.74}$$

by (7.73).

We now follow the argument presented after Theorem (2.45) to obtain as in (2.47) that

$$P_p(\Lambda_r \leq \beta r^{d-1}) \leq \left(\frac{p}{p-p'}\right)^{\beta r^{d-1}} \exp\{-\rho K|\mathcal{K}|\} \quad \text{for } \beta > 0. \tag{7.75}$$

Now,

$$|\mathcal{K}| \geq \nu_1 K^{d-2} \geq \nu_2 \left(\frac{r}{N}\right)^{d-2}$$

by (7.71), where the ν_i are strictly positive constants. Inequality (7.69) follows from (7.75) for suitable small positive β.

We assumed above that r is a multiple of $K+1$. If r is not such a multiple, then one may work instead on $B(R)$ where R is the greatest such multiple satisfying $R \leq r$, and use a 'bounded' amount of probability in order to span the gap $B(r) \setminus B(R)$. □

There follow the proofs of Theorems (7.61) and (7.65).

Proof of Theorem (7.61). It is possible to prove this using the argument used for Theorem (7.2). Instead, we shall use the conclusion of the latter theorem. A different proof is required depending on whether or not $d \geq 3$, and *we assume for the moment that* $d \geq 3$. There are two preliminary lemmas.

For positive integers n and L, we define the subsets $S_n(L)$ and $T_n(L)$ of $\mathbb{Z}^d$ by

$$S_n(L) = [-n, n]^2 \times [0, L]^{d-2}, \tag{7.76}$$

$$T_n(L) = [-n, n]^{d-1} \times [0, L]. \tag{7.77}$$

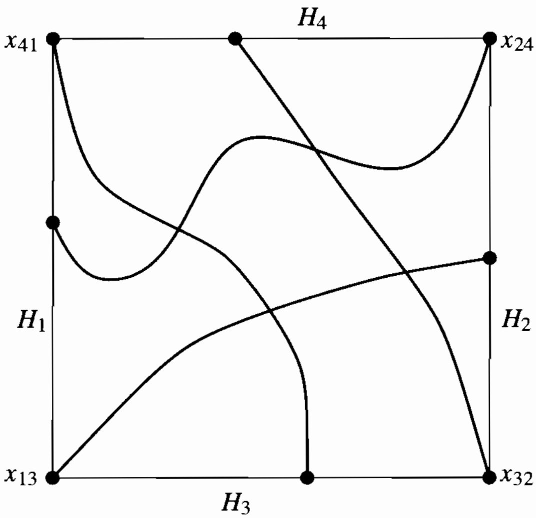


Figure 7.14. A plan of the d-dimensional box $U_m(L)$, showing its 'faces' H_i, together with paths which contribute to the events A_{ij}. Wherever two such paths cross in this plan, there exists a set of at most $(d-2)L$ edges such that the paths would indeed intersect if these edges were re-designated as open.

(7.78) Lemma. *Let $d \geq 3$ and $p > p_c$. There exist a positive integer L and a strictly positive constant $\delta = \delta(p, L)$ such that the following two statements hold:*

(7.79) $\quad P_p\big(x \leftrightarrow y \text{ in } S_n(L)\big) \geq \delta$ *for all $x, y \in S_n(L)$ and $n \geq 1$,*

(7.80) $\quad P_p\big(x \leftrightarrow y \text{ in } T_n(L)\big) \geq \delta$ *for all $x, y \in T_n(L)$ and $n \geq 1$.*

Proof. Since $p > p_c$, we may by Theorem (7.2) find L such that

$$p > p_c\big(\mathbb{Z}_+^2 \times [0, L]^{d-2}\big).$$

Let p' satisfy

(7.81) $$p > p' > p_c\big(\mathbb{Z}_+^2 \times [0, L]^{d-2}\big),$$

and write

$$\theta = P_{p'}\big(0 \leftrightarrow \infty \text{ in } \mathbb{Z}_+^2 \times [0, L]^{d-2}\big),$$

noting that $\theta > 0$.

Let

$$H_1(m) = \{0\} \times [0, m] \times [0, L]^{d-2},$$
$$H_2(m) = \{m\} \times [0, m] \times [0, L]^{d-2},$$
$$H_3(m) = [0, m] \times \{0\} \times [0, L]^{d-2},$$
$$H_4(m) = [0, m] \times \{m\} \times [0, L]^{d-2},$$

be four of the 'faces' of the box $U_m(L) = [0, m]^2 \times [0, L]^{d-2}$, and see Figure 7.14 for an illustration of the argument which follows. Certain pairs of the faces intersect at certain corners of $U_m(L)$, and we shall concentrate on the corners

$$\begin{aligned} x_{13} &= (0, 0, 0, \dots, 0), \\ x_{32} &= (m, 0, 0, \dots, 0), \\ x_{24} &= (m, m, 0, \dots, 0), \\ x_{41} &= (0, m, 0, \dots, 0). \end{aligned}$$

We have that

$$\begin{aligned} \theta &\le P_{p'}\big(0 \leftrightarrow H_2(m) \cup H_4(m) \text{ in } U_m(L)\big) \\ &\le P_{p'}\big(0 \leftrightarrow H_2(m) \text{ in } U_m(L)\big) + P_{p'}\big(0 \leftrightarrow H_4(m) \text{ in } U_m(L)\big), \end{aligned}$$

whence, by symmetry,

$$P_{p'}\big(0 \leftrightarrow H_2(m) \text{ in } U_m(L)\big) \ge \tfrac{1}{2}\theta.$$

Using the rotation invariance of $\mathbb{L}^d$, each of the four increasing events

$$\begin{aligned} A_{13} &= \{x_{13} \leftrightarrow H_2(m) \text{ in } U_m(L)\}, \\ A_{32} &= \{x_{32} \leftrightarrow H_4(m) \text{ in } U_m(L)\}, \\ A_{24} &= \{x_{24} \leftrightarrow H_1(m) \text{ in } U_m(L)\}, \\ A_{41} &= \{x_{41} \leftrightarrow H_3(m) \text{ in } U_m(L)\}, \end{aligned}$$

has probability at least $\frac{1}{2}\theta$. By the FKG inequality,

$$\text{(7.82)} \qquad P_{p'}(A_{13} \cap A_{32} \cap A_{24} \cap A_{41}) \ge (\tfrac{1}{2}\theta)^4.$$

The paths relevant to the events A_{ij} may not intersect one another. However, it is clear from Figure 7.14 that, when all these events occur, there exists a set E containing at most $4(d-2)L$ edges with the property that: if every closed edge in E were to be re-designated open, it would follow that $x_{13} \leftrightarrow x_{32} \leftrightarrow x_{24} \leftrightarrow x_{41}$ in $U_m(L)$. We deduce from (2.49) and (7.81)–(7.82) that there exists $\delta_1 = \delta_1(p, L) > 0$ such that

$$\text{(7.83)} \qquad P_p\big(x_{13} \leftrightarrow x_{32} \leftrightarrow x_{24} \leftrightarrow x_{41} \text{ in } U_m(L)\big) \ge \delta_1 \quad \text{for all } m \ge 1.$$

We shall show the existence of δ_2 such that

$$\text{(7.84)} \qquad P_p\big(z \leftrightarrow 0 \text{ in } S_n(L)\big) \ge \delta_2 > 0 \quad \text{for all } z \in S_n(L),$$

and it will follow by the FKG inequality that

$$\begin{aligned} P_p\big(x \leftrightarrow y \text{ in } S_n(L)\big) &\ge P_p\big(x \leftrightarrow 0 \text{ in } S_n(L)\big) P_p\big(y \leftrightarrow 0 \text{ in } S_n(L)\big) \\ &\ge \delta_2^2 \qquad \text{for all } x, y \in S_n(L), \end{aligned}$$

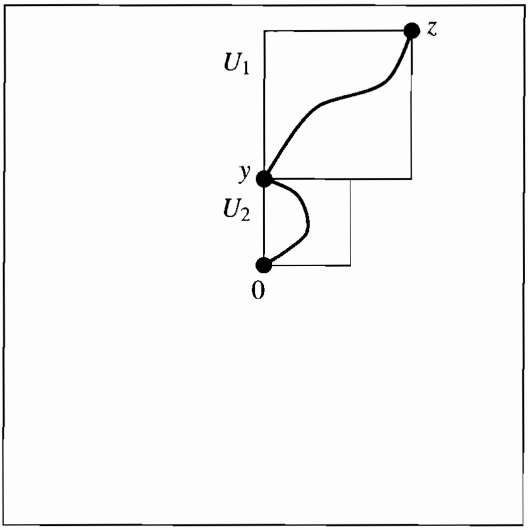


Figure 7.15. An illustration of inequality (7.85).

as required for (7.79). We turn to the proof of (7.84). Let $z = (z_1, z_2, \ldots, z_d) \in S_n(L)$, and assume that $0 \le z_1 \le z_2$; a similar argument is valid otherwise. Define the boxes

$$U_1 = [0, z_1] \times [z_2 - z_1, z_2] \times [0, L]^{d-2},$$
$$U_2 = [0, z_2 - z_1]^2 \times [0, L]^{d-2},$$

and refer to Figure 7.15. We have by the FKG inequality that

$$(7.85) \qquad P_p\big(z \leftrightarrow 0 \text{ in } S_n(L)\big) \ge P_p(z \leftrightarrow y \text{ in } U_1) P_p(y \leftrightarrow 0 \text{ in } U_2)$$

where $y = (0, z_2 - z_1, 0, 0, \ldots, 0)$. Also,

$$(7.86) \qquad P_p(z \leftrightarrow y \text{ in } U_1) \ge p^{(d-2)L} P_p(z' \leftrightarrow y \text{ in } U_1)$$

where $z' = (z_1, z_2, 0, 0, \ldots, 0)$, since there exists a path of U_1 from z to z' having no more than $(d-2)L$ edges. Applying (7.83), we deduce from (7.85)–(7.86) that

$$P_p\big(z \leftrightarrow 0 \text{ in } S_n(L)\big) \ge p^{(d-2)L}\delta_1^2$$

as required for (7.84). This implies (7.79).

We show (7.80) next. Suppose for the moment that $n \ge L$, and let $x = (x_1, x_2, \ldots, x_d) \in T_n(L)$ be such that $x_i \le 0$ for $i = 1, 2, \ldots, d-1$; a similar argument will be valid for general $x \in T_n(L)$. We construct a sequence $s(j)$ of members of $T_n(L)$ as follows:

$$\begin{aligned} s(0) &= x, \\ s(1) &= (0, x_2, x_3, \dots, x_d), \\ s(2) &= (0, 0, x_3, \dots, x_d), \\ &\vdots \\ s(d-3) &= (0, 0, 0, \dots, 0, x_{d-2}, x_{d-1}, x_d), \\ s(d-2) &= (0, 0, 0, \dots, 0). \end{aligned}$$

Let $p > p_c$, and let L, δ be such that (7.79) holds. We claim that

(7.87) $$P_p\big(s(j) \leftrightarrow s(j+1) \text{ in } T_n(L)\big) \geq \delta \quad \text{for } 0 \leq j < d-2,$$

and we prove this as follows. It is the case that, for $0 \leq j < d-2$,

$$s(j), s(j+1) \in (0, 0, \dots, 0, x_{j+3}, x_{j+4}, \dots, x_{d-1}, 0) \\ + [0, L]^j \times [-n, n]^2 \times [0, L]^{d-j-2}.$$

Furthermore, the region on the right is a subset of $T_n(L)$, since $x_i \leq 0$ for $i \leq d-1$. Inequality (7.87) now follows from (7.79).

Applying the FKG inequality, we deduce from (7.87) that

(7.88) $$P_p\big(x \leftrightarrow 0 \text{ in } T_n(L)\big) \geq \delta^{d-2}.$$

Two comments are in order. First, although we have proved (7.88) under the assumption that $x_i \leq 0$ for $i \leq d-1$, the proof is easily adapted to general $x \in T_n(L)$, and we omit the details of this. Secondly, the assumption that $n \geq L$ may be dropped at the expense of replacing δ in (7.88) by some other strictly positive constant δ'. We shall not do this explicitly, but shall accept that (7.88) holds for some $\delta = \delta(p, L) > 0$, and for all $x \in T_n(L)$ and all $n \geq 1$.

Inequality (7.88) implies as below (7.84) that

$$P_p\big(x \leftrightarrow y \text{ in } T_n(L)\big) \geq \delta^{2(d-2)} \quad \text{for all } x, y \in T_n(L).$$

We have proved (7.79) and (7.80) with different constants on the right sides. We now replace these constants by their minimum, thereby obtaining the inequalities as given in the lemma but with a different value of δ. □

In the following we shall sometimes use reals where integers are required. This is for notational convenience only and may easily be rectified.

The next lemma has several applications in its own right. Let n be a positive integer and $\alpha > 1$. For $x, y \in B(n)$, we denote by $E_{n,\alpha}(x, y)$ the event that $x \leftrightarrow \partial B(\alpha n)$, $y \leftrightarrow \partial B(\alpha n)$, but $x \not\leftrightarrow y$ in $B(\alpha n)$; this event is illustrated in Figure 7.16.

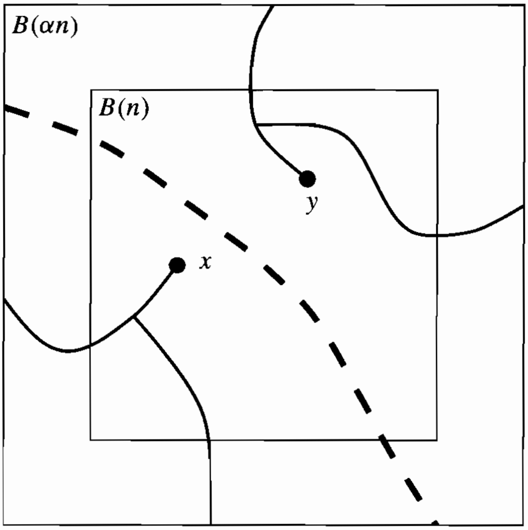


Figure 7.16. An illustration of the event $E_{n,\alpha}(x, y)$.

(7.89) Lemma. *Let $d \geq 3$ and $p > p_c$. There exists $\xi = \xi(p) > 0$ such that*

$$P_p\big(E_{n,\alpha}(x, y)\big) \leq e^{-n(\alpha-1)\xi} \quad \textit{for } x, y \in B(n),\ n \geq 1,\ \alpha > 1.$$

Proof. We aim to show that, with probability at most $e^{-n(\alpha-1)\xi}$, there exist open paths from x and y which intersect $\partial B(\alpha n)$ but which are not connected to one another within $B(\alpha n)$. This we shall do by progressive 'peeling' of skin from $B(\alpha n)$, rather in the manner of the peeling of an orange.

Let L and δ be as in Lemma (7.78), and let $M = L + 1$. We assume for simplicity that αn is an integer, and we express αn in the form

(7.90) $$\alpha n = n + KM + r \quad \text{where} \quad 0 \leq r < M,$$

for some non-negative integer K. For distinct vertices x, y of $B(n)$, we define the event

(7.91) $$A_k(x, y) = \big\{x, y \leftrightarrow \partial B(n + kM),\ x \not\leftrightarrow y \text{ in } B(n + kM - 1)\big\},$$

Now,

(7.92) $$E_{n,\alpha}(x, y) \subseteq A_K(x, y) \subseteq A_{K-1}(x, y) \subseteq \cdots \subseteq A_1(x, y),$$

and therefore,

(7.93) $$P_p\big(E_{n,\alpha}(x, y)\big) \leq \prod_{k=1}^{K} P_p\big(A_k(x, y) \,\big|\, A_{k-1}(x, y)\big).$$

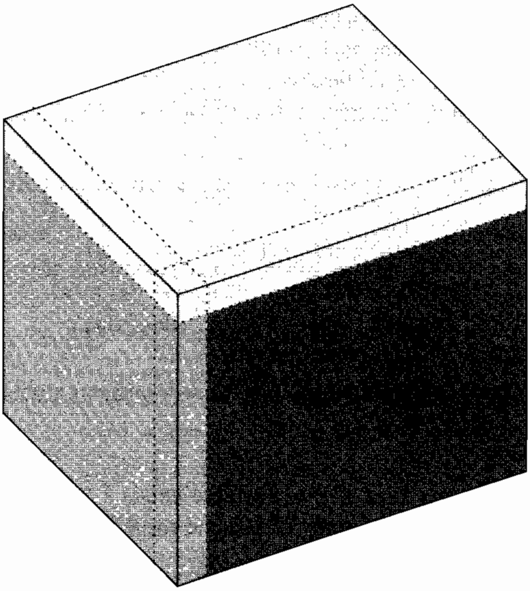


Figure 7.17. The region D_k is the set-theoretic difference of two boxes. It may be expressed as the union of $2d$ 'slices', each being isomorphic to the set $[-r_k, r_k]^{d-1} \times [0, L]$.

We claim that

$$(7.94)\qquad P_p\big(A_k(x, y) \,\big|\, A_{k-1}(x, y)\big) \le 1 - \delta^{d+2} \quad \text{for } k \ge 1.$$

This implies by (7.93) that

$$P_p\big(E_{n,\alpha}(x, y)\big) \le (1 - \delta^{d+2})^K,$$

and the claim of the lemma follows for appropriate ξ. We turn now to the proof of (7.94).

For $z \in B(n)$ and $k \ge 0$, let $V_k(z)$ be the set of all $u \in \partial B(n + kM)$ which are joined to z by open paths all of whose edges use at least one vertex of $B(n+kM-1)$. We note that, on the event $A_k(x, y)$, the sets $V_k(x)$ and $V_k(y)$ are non-empty and disjoint.

Let $k \ge 1$. Given that the event $A_{k-1}(x, y)$ occurs, then $A_k(x, y)$ occurs only if the following holds: $u \nleftrightarrow v$ in the region

$$D_k = B(n + kM - 1) \setminus B(n + (k-1)M - 1),$$

for all $u \in V_{k-1}(x)$, $v \in V_{k-1}(y)$. Now, $V_{k-1}(x)$ and $V_{k-1}(y)$ depend only on the states of edges having at least one endvertex in $B(n + (k-1)M - 1)$, whence

$$P_p\big(A_k(x, y) \,\big|\, A_{k-1}(x, y)\big) \le \sup\big\{P_p(u \nleftrightarrow v \text{ in } D_k) : u, v \in \partial B(n+(k-1)M)\big\}.$$

We shall use (7.80) to prove that

$$(7.95)\qquad P_p(u \leftrightarrow v \text{ in } D_k) \ge \delta^{d+2} \quad \text{for all } u, v \in D_k,$$

which implies (7.94).

The region D_k may be thought of as a type of d-dimensional 'shell', in rough terms comprising $2d$ overlapping 'slices' each being isomorphic to the region $T_{r_k}(L) = [-r_k, r_k]^{d-1} \times [0, L]$ where $r_k = n + kM - 1$. We do not say exactly what these $2d$ slices are, instead illustrating in Figure 7.17 the situation when $d = 3$.

Let $z(1), z(2), \dots, z(d+1)$ be a collection of 'corners' of $B(r_k)$, given by

$$\begin{aligned} z(1) &= (-r_k, -r_k, \dots, -r_k, -r_k), \\ z(2) &= (r_k, -r_k, -r_k, \dots, -r_k, -r_k), \\ z(3) &= (r_k, r_k, -r_k, \dots, -r_k, -r_k), \\ &\vdots \\ z(d+1) &= (r_k, r_k, \dots, r_k, r_k). \end{aligned}$$

For each consecutive pair $z(j), z(j+1)$, there exists a 'slice' of D_k containing both $z(j)$ and $z(j+1)$. It follows by (7.80) and the FKG inequality that

$$\begin{aligned} P_p\big(z(1) \leftrightarrow z(2) \leftrightarrow \cdots \leftrightarrow z(d+1) \text{ in } D_k\big) &\geq \prod_{i=1}^{d} P_p\big(z(i) \leftrightarrow z(i+1) \text{ in } D_k\big) \\ &\geq \delta^d. \end{aligned}$$

Let $u, v \in \partial B(n + (k-1)M)$. There exists a 'slice' of D_k containing both u and some $z(i)$, and similarly there exists a slice containing both v and some $z(j)$. By (7.80) and the FKG inequality,

$$P_p(u \leftrightarrow v \text{ in } D_k) \geq P_p\big(u \leftrightarrow z(i) \leftrightarrow z(j) \leftrightarrow v \text{ in } D_k\big) \geq \delta^{d+2}, \tag{7.96}$$

and (7.95) has been proved. □

We shall make use of Lemma (7.89) in our proof of the next lemma.

(7.97) Lemma. *Let $d \geq 3$, $0 < \varepsilon < 1$, and $p > p_c$. It is the case that*

(7.98)

$$P_p\Big(B(n) \text{ has a crossing cluster } C \text{ satisfying } |C| \geq (1-\varepsilon)\theta(p)|B(n)|\Big) \to 1 \qquad \textit{as } n \to \infty.$$

Proof. Assume that $d \geq 3$, $0 < \varepsilon < 1$, and $p > p_c$. Let ξ be given as in Lemma (7.89), and choose ν such that $\nu\xi > 2d$. Let K_n be the set of all vertices of $B(n - \nu \log n)$ which belong to infinite open clusters of $\mathbb{L}^d$, and let I_n be the open subgraph of $B(n)$ having vertex set $\{x : x \leftrightarrow K_n \text{ in } B(n)\}$ together with all open edges joining pairs of such vertices. We claim that I_n is (with large probability) a connected open cluster of $B(n)$. Suppose that, on the contrary, I_n is disconnected. Then there exist $u, v \in K_n$ such that $u \nleftrightarrow v$ in $B(n)$. By Lemma (7.89),

$$\text{(7.99)}\quad P_p(I_n \text{ is disconnected}) \le |B(n)|^2 \exp\left\{-(n - \nu \log n)\frac{\nu \log n}{n}\xi\right\}$$
$$\le \frac{A}{n^{\nu\xi - 2d}} \to 0 \quad \text{as } n \to \infty,$$

for some constant A.

Next we study the cardinality of I_n. We have that

$$\text{(7.100)}\qquad |I_n| \ge |K_n| = \sum_{x \in B(n - \nu \log n)} I_{\{x \leftrightarrow \infty\}},$$

the sum of indicator random variables each having mean value $\theta(p)$. By the ergodic theorem,

$$\text{(7.101)}\qquad P_p\Big(|K_n| \ge (1 - \tfrac{1}{2}\varepsilon)\theta(p)|B(n - \nu \log n)|\Big) \to 1 \quad \text{as } n \to \infty.$$

Since $|B(n - \nu \log n)|/|B(n)| \to 1$ as $n \to \infty$, we deduce that

$$\text{(7.102)}\qquad P_p\big(|I_n| \ge (1 - \varepsilon)\theta(p)|B(n)|\big) \to 1 \quad \text{as } n \to \infty.$$

It follows from (7.99) and (7.102) that, with probability approaching 1 as $n \to \infty$, I_n is a connected component of $B(n)$ having cardinality at least $(1-\varepsilon)\theta(p)|B(n)|$.

Next we show that

$$\text{(7.103)}\qquad P_p\big(I_n \text{ is a crossing cluster of } B(n)\big) \to 1 \quad \text{as } n \to \infty.$$

For $i \in \{1, 2, \dots, d\}$ and $\sigma \in \{-, +\}$, we define the face $F_{i,\sigma}$ of $B(n)$ by

$$F_{i,\sigma} = \big\{x \in B(n) : x_i = \sigma n\big\}.$$

Let

$$E_{i,\sigma} = \{I_n \cap F_{i,\sigma} \ne \varnothing\}.$$

By symmetry and the FKG inequality,

$$P_p(K_n = \varnothing) \ge P_p\bigg(\bigcap_{i,\sigma} \overline{E_{i,\sigma}}\bigg) \ge \Big[P_p(\overline{E_{1,+}})\Big]^{2d}.$$

It follows by the FKG inequality and (7.101) that

$$P_p\bigg(\bigcap_{i,\sigma} E_{i,\sigma}\bigg) \ge \prod_{i,\sigma} P_p(E_{i,\sigma}) = \big\{1 - P_p(\overline{E_{1,+}})\big\}^{2d}$$
$$\ge \left\{1 - \Big[P_p(K_n = \varnothing)\Big]^{1/(2d)}\right\}^{2d}$$
$$\to 1 \quad \text{as } n \to \infty.$$

If every $E_{i,\sigma}$ occurs, and in addition I_n is connected, then I_n is a crossing cluster. Combined with (7.99) and (7.102), this implies (7.98) with $C = I_n$. □

Let $T_{m,n}$ be the event that $B(n)$ has a crossing cluster C and, in addition, $B(n)$ contains some *other* open cluster D having diameter at least m.

(7.104) Lemma. *Let $d \ge 3$ and $p > p_c$. There exists $\mu = \mu(p) > 0$ such that*

$$P_p(T_{m,n}) \le d(2n+1)^{2d} e^{-\mu m} \quad \textit{for } m, n \ge 1.$$

Proof. This resembles the proof of Lemma (7.89). Let

$$\begin{aligned} H_r &= \{x \in B(n) : x_1 = r\} \\ &= \{r\} \times [-n, n]^{d-1} \quad \text{for } -n \le r \le n, \end{aligned}$$

and let

$$U_{i,j} = [i, j) \times [-n, n]^{d-1} \quad \text{for } -n \le i < j \le n.$$

For distinct vertices x, y of $B(n)$ with $x_1 = y_1 = \varphi$ say, we define the event

$$A_k(x, y) = \left\{x, y \leftrightarrow H_{\varphi+k} \text{ in } B(n), x \not\leftrightarrow y \text{ in } U_{\varphi,\varphi+k}\right\}.$$

Suppose now that $T_{m,n}$ occurs for some m, n; let C be a crossing cluster, and let D be another open cluster of $B(n)$ with $\text{diam}(D) \ge m$. Let us assume for the moment that the diameter of D is achieved in the first coordinate direction, which is to say that:

(7.105) $\qquad$ there exist $u, v \in D$ with $v_1 - u_1 = \text{diam}(D) \ge m$.

If this occurs, there must exist an integer φ satisfying

(7.106) $$-n \le \varphi \le n - m$$

together with two vertices $x, y \in H_\varphi$ with the property that the event $A_m(x, y)$ occurs. Therefore, using the symmetry of $\mathbb{L}^d$ and the fact that there are d coordinate directions,

(7.107) $$\begin{aligned} P_p(T_{m,n}) &\le d \sum_{\substack{\varphi: \\ -n \le \varphi \le n-m}} \sum_{x,y \in H_\varphi} P_p\big(A_m(x, y)\big) \\ &\le d|B(n)|^2 \sup_{x,y \in H_{-n}} \left\{P_p\big(A_m(x, y)\big)\right\}. \end{aligned}$$

We propose to show next that the last probability is small, uniformly in the choice of x and y.

Suppose that $p > p_c$, choose L and δ according to Lemma (7.78), and set $M = L + 1$. For $1 \le m \le 2n$, we may express m in the form

(7.108) $$m = KM + r \quad \text{where} \quad 0 \le r < M,$$

and K is a non-negative integer. For $x, y \in H_{-n}$, we have that

$$A_m(x, y) \subseteq A_{KM}(x, y) \subseteq A_{(K-1)M}(x, y) \subseteq \cdots \subseteq A_M(x, y),$$

whence

(7.109) $$P_p\big(A_m(x,y)\big) \le \prod_{i=1}^{K} P_p(A_{iM} \mid A_{(i-1)M}).$$

For $z \in H_{-n}$ and $i \ge 0$, let $V_i(z)$ denote the set of all vertices $u \in H_{-n+iM}$ which are joined to z by open paths of $B(n)$ whose intersection with H_{-n+iM} is the singleton u. The set $V_i(z)$ depends on the states of edges having at least one endvertex to the left of H_{-n+iM} (that is, having at least one endvertex w with $w_1 < -n + iM$). Note that, for distinct vertices x, y of H_{-n}, the sets $V_i(x)$ and $V_i(y)$ are non-empty and disjoint on the event $A_{iM}(x, y)$.

Let $i \ge 1$. Conditional on the event $A_{(i-1)M}(x,y)$, the event $A_{iM}(x,y)$ occurs only if the following holds: $u \nleftrightarrow v$ in $U_{-n+(i-1)M,-n+iM}$ for all $u \in V_{i-1}(x)$ and $v \in V_{i-1}(y)$. Now $U_{-n+(i-1)M,-n+iM}$ is isomorphic to $[-n,n]^{d-1} \times [0, L]$, whence by (7.80) there exists $\delta = \delta(p, L) > 0$ such that

$$\begin{aligned} &P_p\big(A_{iM}(x,y) \,\big|\, A_{(i-1)M}(x,y)\big) \\ &\qquad \le \sup\Big\{P_p\big(u \nleftrightarrow v \text{ in } U_{-n+(i-1)M,-n+iM}\big) : u, v \in H_{-n+(i-1)M}\Big\} \\ &\qquad \le 1 - \delta. \end{aligned}$$

We conclude by (7.108) and (7.109) that

$$P_p\big(A_m(x,y)\big) \le (1-\delta)^K \le (1-\delta)^{(m/M)-1}.$$

The claim of the lemma follows by (7.107). □

We shall now complete the proof of Theorem (7.61) when $d \ge 3$. Consider the event that $B(n)$ contains a crossing cluster C satisfying $|C| \ge (1-\varepsilon)\theta(p)|B(n)|$, and in addition there exists some other open cluster D of $B(n)$ with $\text{diam}(D) \ge n$. By Lemma (7.104), this event has probability not exceeding $Ae^{-\mu n}$ for some constants $A = A(p)$ and $\mu = \mu(p) > 0$. It follows by Lemma (7.97) that

$$P_p\big(B(n) \text{ is } \varepsilon\text{-good}\big) \to 1 \quad \text{as } n \to \infty$$

for all ε satisfying $0 < \varepsilon < 1$, and the proof of (7.62) is complete in this case.

Finally we consider the case $d = 2$. We sketch the details of this, and shall rely on results to be proved in Chapter 11. Assume that $p > p_c$.

Let $m, n \ge 1$, and let $L_{m,n}$ be the event that the rectangle $[0, m] \times [-n, n]$ contains an open crossing joining its two sides of length m. Adapting the argument prior to Lemma (11.22), we have that

(7.110) $$P_p(L_{m,n}) \ge 1 - Ane^{-\sigma m}$$

for some strictly positive constants A, σ.

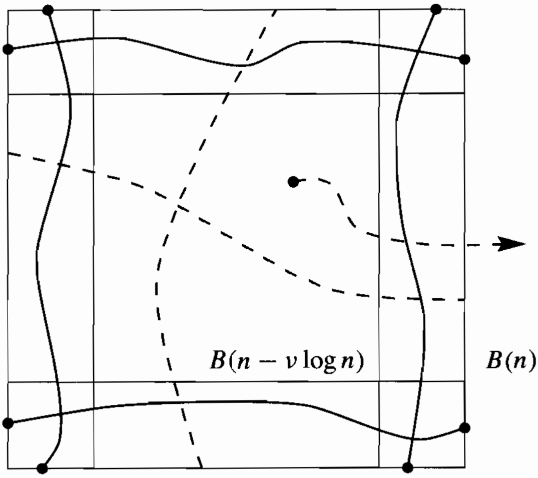


Figure 7.18. The outer tubes are crossed in their long directions by open paths, and the interior region contains a large number of vertices lying in infinite open paths. All such vertices are joined to one another by open paths of $B(n)$, and the ensuing cluster is a crossing cluster.

With $\nu > \sigma^{-1}$, we deduce in particular that

(7.111) $$P_p(L_{\nu \log n, n}) \to 1 \quad \text{as } n \to \infty.$$

Let O_n be the event that each of the rectangles

$$[-n, n] \times [n - \nu \log n, n],$$
$$[-n, n] \times [-n, -n + \nu \log n],$$
$$[-n, -n + \nu \log n] \times [-n, n],$$
$$[n - \nu \log n, n] \times [-n, n],$$

is crossed in the 'long direction' by an open path. This event is illustrated in Figure 7.18. Using (7.111), the symmetry of $\mathbb{L}^2$, and the FKG inequality,

(7.112) $$P_p(O_n) \to 1 \quad \text{as } n \to \infty.$$

Let $\varepsilon > 0$, and let D_n be the event that the box $B(n - \nu \log n)$ contains at least $(1 - \varepsilon)\theta(p)|B(n)|$ vertices which lie in infinite open paths. We have by the ergodic theorem as in (7.101)–(7.102) that

$$P_p(D_n) \to 1 \quad \text{as } n \to \infty.$$

If both D_n and O_n occur, then all such vertices lie in the same crossing cluster C of $B(n)$ satisfying:

(a) $|C| \geq (1 - \varepsilon)\theta(p)|B(n)|$, and

(b) C contains all top–bottom and left–right crossings of $B(n)$.

It remains to eliminate the possibility that $B(n)$ may contain some other open cluster having diameter at least n. Such a cluster must traverse between the long sides of some rectangle having the form either $[-n,n]\times[i,i+n]$ or $[i,i+n]\times[-n,n]$ where $-n\le i\le 0$. By (7.110), every such rectangle is traversed by an open path between its sides of length n, with probability at least

$$1-2An(n+1)e^{-\sigma n}\to 1 \quad \text{as } n\to\infty. \tag{7.113}$$

If this last event occurs, and in addition $O_n\cap D_n$ occurs, then there can exist no open cluster other than C having diameter at least n. The claim of the theorem follows when $d=2$. □

Proof of Theorem (7.65). We follow Liggett, Schonmann, and Stacey (1997).

Suppose that Y is a k-dependent family satisfying (7.66) with $0\le\delta\le 1$. The idea is to 'dilute' the Y_x a little, in order to obtain the required stochastic domination. To this end, assume that there exist $\alpha,\rho\in(0,1)$ such that

$$(1-\alpha)(1-\rho)^{|B(k)|}\ge 1-\delta, \tag{7.114}$$

$$(1-\alpha)\alpha^{|B(k)|}\ge 1-\delta, \tag{7.115}$$

and let $Z^\rho=\{Z_x^\rho : x\in\mathbb{Z}^d\}$ be a family of independent random variables taking values in the set $\{0,1\}$, independent of the family Y, and satisfying

$$P(Z_x^\rho=1)=1-P(Z_x^\rho=0)=\rho.$$

We claim that the family $Z^\rho Y=\{Z_x^\rho Y_x : x\in\mathbb{Z}^d\}$ satisfies

$$Z^\rho Y\ge_{\text{st}} Z^{\alpha\rho}. \tag{7.116}$$

It is trivial from (7.63) that $Y\ge_{\text{st}} Z^\rho Y$, and therefore (7.116) implies that $Y\ge_{\text{st}} Z^{\alpha\rho}$. Now, as δ approaches 1, we may allow α and ρ to approach 1 also, whence we may find $\alpha=\alpha(\delta)$ and $\rho=\rho(\delta)$ satisfying (7.114)–(7.115) such that

$$\pi(\delta)=\alpha(\delta)\rho(\delta)\to 1 \quad \text{as } \delta\to 1.$$

This will imply the claim of the theorem.

In order to obtain (7.116), we shall prove the following statement by induction. Let $j\ge 0$, and let $x(1),x(2),\dots,x(j+1)$ be distinct points in $\mathbb{Z}^d$. It is the case that

$$P\big(Y_{x(j+1)}=1 \,\big|\, Z_{x(i)}^\rho Y_{x(i)}=z_i \text{ for } 1\le i\le j\big)\ge\alpha \tag{7.117}$$

for all $z_1,z_2,\dots,z_j$ for which

$$P\big(Z_{x(i)}^\rho Y_{x(i)}=z_i \text{ for } 1\le i\le j\big)>0. \tag{7.118}$$

Consider first the case $j = 0$. We have that

$$P(Y_x = 1) \geq \delta \quad \text{for } x \in \mathbb{Z}^d,$$

by (7.66). Also, $\alpha \leq \delta$ by (7.115), whence the claim holds with $j = 0$.

Suppose now that the claim holds whenever $j < J$ where $J \geq 1$, and set $j = J$. Let $z_1, z_2, \ldots, z_J$ satisfy (7.118), and partition $\{x(1), x(2), \ldots, x(J)\}$ as $N^0 \cup N^1 \cup M$ where

$$\begin{aligned} N^0 &= \big\{x(i) : 1 \leq i \leq J,\ \delta(x(i), x(J+1)) \leq k,\ z_i = 0\big\}, \\ N^1 &= \big\{x(i) : 1 \leq i \leq J,\ \delta(x(i), x(J+1)) \leq k,\ z_i = 1\big\}, \\ M &= \{x(1), x(2), \ldots, x(J)\} \backslash (N^0 \cup N^1). \end{aligned}$$

By the properties of Z^ρ,

$$\begin{aligned} P\big(Y_{x(J+1)} = 1 \,\big|\, Z^\rho_{x(i)} Y_{x(i)} = z_i \text{ for } 1 \leq i \leq J\big) \\ = P\big(Y_{x(J+1)} = 1 \,\big|\, A^0 \cap A^1 \cap A\big) \end{aligned}$$

where

$$\begin{aligned} A^0 &= \{Z^\rho_{x(i)} Y_{x(i)} = 0 \text{ for } i \in N^0\}, \\ A^1 &= \{Y_{x(i)} = 1 \text{ for } i \in N^1\}, \\ A &= \{Z^\rho_{x(i)} Y_{x(i)} = z_i \text{ for } i \in M\}. \end{aligned}$$

Now,

(7.119)

$$\begin{aligned} P\big(Y_{x(J+1)} = 1 \,\big|\, A^0 \cap A^1 \cap A\big) &= 1 - \frac{P(Y_{x(J+1)} = 0,\ A^0 \cap A^1 \cap A)}{P(A^0 \cap A^1 \cap A)} \\ &\geq 1 - \frac{P(Y_{x(J+1)} = 0,\ A)}{P(B^0 \cap A^1 \cap A)} \\ &= 1 - \frac{P(Y_{x(J+1)} = 0 \mid A)}{P(B^0 \cap A^1 \mid A)} \end{aligned}$$

where

$$B^0 = \{Z^\rho_{x(i)} = 0 \text{ for } i \in N^0\} \subseteq A^0.$$

Since Y is k-dependent, and since M contains no vertex within distance k of $x(J+1)$, we have that

(7.120) $$P\big(Y_{x(J+1)} = 0 \,\big|\, A\big) = P\big(Y_{x(J+1)} = 0\big) \leq 1 - \delta$$

by (7.66). In addition,

(7.121) $$P(B^0 \cap A^1 \mid A) = (1 - \rho)^{|N^0|} P(A^1 \mid A),$$

and it remains to bound $P(A^1 \mid A)$ from below, which we do by way of the induction hypothesis.

Assume that $N^1 = \{y_1, y_2, \ldots, y_n\}$ where $n \geq 1$. We have that

$$P(A^1 \mid A) = \prod_{k=1}^{n} P\big(Y_{y_k} = 1 \,\big|\, A,\ Y_{y_i} = 1 \text{ for } 1 \leq i < k\big).$$

Furthermore, using the properties of Z^ρ and the induction hypothesis,

$$\begin{aligned} P\big(Y_{y_k} = 1 \,\big|\, A,\ Y_{y_i} = 1 \text{ for } 1 \leq i < k\big) \\ &= P\big(Y_{y_k} = 1 \,\big|\, A,\ Z^\rho_{y_i} Y_{y_i} = 1 \text{ for } 1 \leq i < k\big) \\ &\geq \alpha, \end{aligned}$$

and therefore,

$$P(A^1 \mid A) \geq \alpha^{|N^1|} \quad \text{if } |N^1| \geq 1. \tag{7.122}$$

The same inequality holds trivially when $N^1 = \varnothing$. Substituting (7.120)–(7.122) into (7.119), we deduce that

$$P\big(Y_{x(J+1)} = 1 \,\big|\, A^0 \cap A^1 \cap A\big) \geq 1 - \frac{1-\delta}{(1-\rho)^{|N^0|}\alpha^{|N^1|}}.$$

Now $|N^0| + |N^1| \leq |B(k)|$, whence by (7.114)–(7.115) the right side is at least

$$1 - \frac{1-\delta}{(1-\delta)/(1-\alpha)} = \alpha.$$

This completes the induction step, and we deduce that (7.117) holds for all j and all $z_1, z_2, \ldots, z_j$ satisfying (7.118).

It follows by (7.117) and independence that

$$P_p\big(Z^\rho_{x(j+1)} Y_{x(j+1)} = 1 \,\big|\, Z^\rho_{x(i)} Y_{x(i)} = z_i \text{ for } 1 \leq i \leq j\big) \geq \alpha\rho$$

for all j, whence (7.116) holds by the discussion around (7.64). The proof is complete. □

7.5 Notes

Section 7.1. In Chapter 8, we shall encounter several applications of the statement $p_c^{\text{slab}} = p_c$. Note that the 'slabs' in this book are two-dimensional; in earlier work they were sometimes taken to be $(d-1)$-dimensional. The importance of slab percolation was realized first by Aizenman, Chayes, Chayes, Fröhlich, and Russo (1983).

Section 7.2. Theorem (7.2) is taken from Grimmett and Marstrand (1990). Their ideas are related to those of Barsky, Grimmett, and Newman (1991a, b).

Section 7.3. Percolation in half-spaces was studied by Barsky, Grimmett, and Newman (1991a, b). They used the block arguments of this section to prove that $p_c^{\text{slab}} = p_c(\mathbb{H})$ and $\theta_{\mathbb{H}}(p_c(\mathbb{H})) = 0$ when $d \geq 3$, but their arguments fell short of a proof that $p_c(\mathbb{H}) = p_c$. The conclusion of Theorem (7.35) may be generalized somewhat using essentially the same proof. In the case $d = 2$, earlier results of Harris (1960) and Kesten (1980a) imply that $\theta_{\mathbb{H}}(p_c) = 0$; see Chapter 11. It is perhaps surprising that no simple proof of the equality of $p_c(\mathbb{H})$ and p_c is known when $d \geq 3$.

Section 7.4. Static block arguments have been in use for percolation since Kesten (1981) or earlier. The block variables of this section are based on those used by Pisztora (1996) and Deuschel and Pisztora (1996) to prove large deviation estimates for quantities arising in percolation and random-cluster models. There are various ways of introducing such variables to suit the application in mind.

The value of ε in the definition of the term 'ε-good' is arbitrary. Only slightly more work is needed with ε replaced by some appropriate ε_n satisfying $\varepsilon_n \downarrow 0$ as $n \to \infty$. The necessary work is implicit in papers devoted to central limit theorems for the number of vertices in a box which lie in infinite open paths; see the notes for Section 11.6.

Further applications of static renormalization may be found in Pisztora (1996), Deuschel and Pisztora (1996), Antal and Pisztora (1996), and Bezuidenhout and Grimmett (1997).

The proofs of Theorem (7.68) and Lemma (7.78) use the 'sprinkling' technique introduced by Aizenman, Chayes, Chayes, Fröhlich, and Russo (1983). Lemma (7.89) has its origins in Kesten and Zhang (1990) and Grimmett and Marstrand (1990). The arguments presented here owe much to Kesten and Zhang (1990) and Pisztora (1996).

Comparison inequalities between k-dependent families and product measures have been used by many authors including Durrett (1988), Andjel (1993), and Penrose and Pisztora (1996). The result of Theorem (7.65) was proved in a more general form by Liggett, Schonmann, and Stacey (1997), to which the reader is referred for background material concerning stochastic domination.

Chapter 8

The Supercritical Phase

8.1 Introduction

In this chapter we shall consider bond percolation on $\mathbb{L}^d$ where $d \geq 2$, and we shall suppose that $p > p_c$. In this case, there exists at least one infinite open cluster, and the first main result of the chapter is that there exists a unique such cluster.

In the subcritical phase we were interested in such quantities as $P_p(|C(0)| \geq n)$ and $P_p(0 \leftrightarrow x)$, and particularly in their asymptotic behaviour for large n and $|x|$. Such quantities are of less interest in the supercritical phase, since their asymptotic behaviour is dominated to the first order by the probability that the origin belongs to the infinite open cluster, or the probability that both the origin and x are in this cluster, respectively. Of greater interest are the corresponding probabilities defined in terms of *finite* open clusters. That is to say, we shall study the asymptotic behaviour of $P_p(n \leq |C(0)| < \infty)$ and $P_p(0 \leftrightarrow x,\ |C(0)| < \infty)$.

Here is a guide to the chapter. We begin by proving that, if there exists an infinite open cluster, then there is (almost surely) exactly one such cluster. As a consequence of this, the percolation probability $\theta(p)$ is a continuous function of p on $(p_c, 1]$. There follow later two sections devoted respectively to the exponential decay of (a) the tail of the radius of a finite open cluster, and (b) the truncated connectivity function

$$\tau_p^{\mathrm{f}}(0, x) = P_p\big(0 \leftrightarrow x,\ |C| < \infty\big);$$

we prove that these quantities decay *strictly* exponentially (that is, with *strictly* positive rate) when $p > p_c$. This is followed by a proof that the cluster size distribution decays slower than exponentially in the supercritical phase. In Section 8.7, we prove that the percolation probability $\theta(p)$, the mean size $\chi^{\mathrm{f}}(p)$ of the finite open cluster at the origin, and the number $\kappa(p)$ of open clusters per vertex are infinitely differentiable functions of p when $p > p_c$. Finally, we reach some simple conclusions about the geometry of the infinite open cluster.

8.2 Uniqueness of the Infinite Open Cluster

The principal result of this section is the following: for any value of p for which there is a strictly positive probability of an infinite open cluster, there exists almost surely exactly one such cluster.

(8.1) Theorem. Uniqueness of the infinite open cluster. *If p is such that $\theta(p) > 0$, then*

$$P_p\big(\text{there exists exactly one infinite open cluster}\big) = 1.$$

Proof. We follow Burton and Keane (1989). The claim is trivial if $p = 0, 1$, and we assume henceforth that $0 < p < 1$. Let N be the number of infinite open clusters. Let B be a finite set of vertices, and let $\mathbb{E}_B$ denote the set of edges of $\mathbb{L}^d$ joining pairs of vertices in B. We write $N_B(0)$ (respectively $N_B(1)$) for the number of infinite open clusters when all edges in $\mathbb{E}_B$ are declared to be closed (respectively open). Finally, M_B denotes the number of infinite open clusters which intersect B.

The sample space $\Omega = \{0, 1\}^{\mathbb{E}^d}$ is a product space with a natural family of translations inherited from the translations of the lattice $\mathbb{L}^d$. Furthermore, P_p is a product measure on Ω. Since N is a translation-invariant function on Ω, it is a.s. constant, which is to say that

(8.2) there exists $k \in \{0, 1, 2, \dots\} \cup \{\infty\}$ such that $P_p(N = k) = 1$.

Naturally, the value of k depends on the choice of p.

Next we show that the k in (8.2) necessarily satisfies $k \in \{0, 1, \infty\}$. Suppose on the contrary that (8.2) holds with some k satisfying $2 \le k < \infty$. Let B be a finite set of vertices. Since every configuration on $\mathbb{E}_B$ has a strictly positive probability, it follows by the almost sure constantness of N that

$$P_p\big(N_B(0) = N_B(1) = k\big) = 1.$$

Now $N_B(0) = N_B(1)$ if and only if B intersects at most one infinite open cluster (this is where we use the assumption that $k < \infty$), and therefore

$$P_p(M_B \ge 2) = 0.$$

Clearly, M_B is non-decreasing in B, and $M_B \to N$ as $B \uparrow \mathbb{Z}^d$. We take B to be the 'diamond' $S(n) = \{x \in \mathbb{Z}^d : \delta(0, x) \le n\}$, and take the limit as $n \to \infty$ to find that

$$0 = P_p\big(M_{S(n)} \ge 2\big) \to P_p(N \ge 2), \tag{8.3}$$

which is to say that $k \leq 1$, in contradiction of the assumption $2 \leq k < \infty$.

It remains to rule out the case $k = \infty$. Suppose that $k = \infty$. We will derive a contradiction by using a geometrical argument. We call a vertex x a *trifurcation* if:

(a) x lies in an infinite open cluster;

(b) there exist exactly three open edges incident to x; and

(c) the deletion of x and its three incident open edges splits this infinite cluster into exactly three disjoint infinite clusters and no finite clusters;

and we denote by T_x the event that x is a trifurcation. Now, $P_p(T_x)$ is constant for all x, and therefore

$$\frac{1}{|B(n)|} E_p\left(\sum_{x \in B(n)} I_{T_x}\right) = P_p(T_0). \tag{8.4}$$

(Recall that I_A denotes the indicator function of an event A.) It will be useful to know that the quantity $P_p(T_0)$ is strictly positive, and it is here that we use the assumed infinity of infinite clusters. Let $M_B(0)$ be the number of infinite open clusters which intersect B when all edges of $\mathbb{E}_B$ are declared closed. Since $M_B(0) \geq M_B$, we have by the remarks around (8.3) that

$$P_p\big(M_{S(n)}(0) \geq 3\big) \geq P_p\big(M_{S(n)} \geq 3\big) \to P_p(N \geq 3) = 1 \qquad \text{as } n \to \infty.$$

Therefore, there exists n such that

$$P_p\big(M_{S(n)}(0) \geq 3\big) \geq \tfrac{1}{2},$$

and we set $B = S(n)$. Note that:

(a) the event $\{M_B(0) \geq 3\}$ is independent of the states of edges in $\mathbb{E}_B$;

(b) if the event $\{M_B(0) \geq 3\}$ occurs, there exist x, y, z $(\in \partial B)$ lying in distinct infinite open clusters of $\mathbb{E}^d \setminus \mathbb{E}_B$.

Let $\omega \in \{M_B(0) \geq 3\}$, and pick $x = x(\omega)$, $y = y(\omega)$, $z = z(\omega)$ according to (b). (If there is more than one possible such triple, we pick such a triple according to some predetermined rule.) It is a minor geometrical exercise (see Figure 8.1) to verify that there exist in $\mathbb{E}_B$ three paths joining the origin to (respectively) x, y, and z, and that these paths may be chosen in such a way that:

(i) the origin is the unique vertex common to any two of them; and

(ii) each touches exactly one vertex lying in ∂B.

Let $J_{x,y,z}$ be the event that all the edges in these paths are open, and that all other edges in $\mathbb{E}_B$ are closed.

Since B is finite,

$$P_p\big(J_{x,y,z} \,\big|\, M_B(0) \geq 3\big) \geq \big(\min\{p, 1-p\}\big)^R > 0,$$

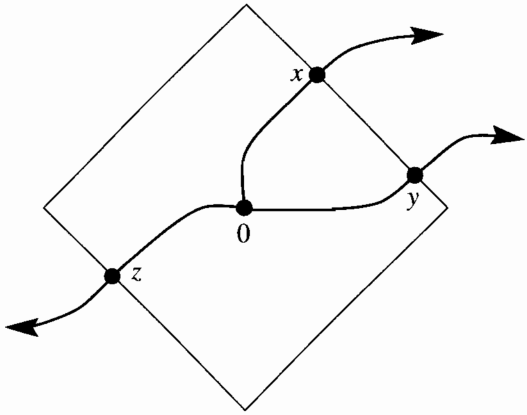


Figure 8.1. Take a diamond B which intersects at least three distinct infinite open clusters, and then alter the configuration inside B in order to create a configuration in which 0 is a trifurcation.

where R is the number of edges in $\mathbb{E}_B$. Now,

$$P_p\big(0 \text{ is a trifurcation}\big) \geq P_p\big(J_{x,y,z} \,\big|\, M_B(0) \geq 3\big) P_p\big(M_B(0) \geq 3\big)$$
$$\geq \tfrac{1}{2}\big(\min\{p, 1-p\}\big)^R > 0,$$

which is to say that $P_p(T_0) > 0$ in (8.4).

It follows from (8.4) that the mean number of trifurcations inside $B(n)$ grows in the manner of $|B(n)|$ as $n \to \infty$. From this we will derive a contradiction, and the following rough argument will convince some. Select a trifurcation (t_1, say) of $B(n)$, and choose some vertex y_1 ($\in \partial B(n)$) which satisfies $t_1 \leftrightarrow y_1$ in $B(n)$. We now select a new trifurcation $t_2 \in B(n)$. It may be seen, using the definition of the term 'trifurcation', that there exists $y_2 \in \partial B(n)$ such that $y_1 \neq y_2$ and $t_2 \leftrightarrow y_2$ in $B(n)$. We continue similarly, at each stage picking a new trifurcation $t_k \in B(n)$ and a new vertex $y_k \in \partial B(n)$. If there exist τ trifurcations in $B(n)$, then we obtain τ distinct vertices y_k lying in $\partial B(n)$. Therefore $|\partial B(n)| \geq \tau$. However, by the remarks above, $E_p(\tau)$ is comparable to $|B(n)|$. This is a contradiction for large n, since $|\partial B(n)|$ grows in the manner of n^{d-1} and $|B(n)|$ grows in the manner of n^d.

There follows a rigorous version of this argument. Let Y be a finite set with $|Y| \geq 3$, and define a *3-partition* $\Pi = \{P_1, P_2, P_3\}$ of Y to be a partition of Y into exactly three non-empty sets P_1, P_2, P_3. Given two 3-partitions $\Pi = \{P_1, P_2, P_3\}$ and $\Pi' = \{P_1', P_2', P_3'\}$, we say that Π and Π' are *compatible* if there exists an ordering of their elements such that $P_1 \supseteq P_2' \cup P_3'$ (or, equivalently, such that $P_1' \supseteq P_2 \cup P_3$). A collection $\mathcal{P}$ of 3-partitions is said to be *compatible* if each distinct pair therein is compatible.

(8.5) Lemma. *If $\mathcal{P}$ is a compatible family of distinct 3-partitions of Y, then $|\mathcal{P}| \leq |Y| - 2$.*

Proof. We prove this by induction on $|Y|$, noting first that $|\mathcal{P}| \le 1$ if $|Y| = 3$. Assume that the claim holds whenever $|Y| \le n$, and let Y satisfy $|Y| = n + 1$. Pick a singleton $y \in Y$, and write $Z = Y \setminus \{y\}$.

Let $\mathcal{P}$ be a family of compatible 3-partitions of Y. Any $\Pi \in \mathcal{P}$ may be expressed in the form $\Pi = \{\Pi_1 \cup \{y\}, \Pi_2, \Pi_3\}$ for some disjoint subsets Π_1, Π_2, Π_3 of Z satisfying: Π_2 and Π_3 are non-empty, and $\Pi_1 \cup \Pi_2 \cup \Pi_3 = Z$. Let $\mathcal{P}'$ be the set of all such Π such that $\Pi_1 \neq \varnothing$, and let $\mathcal{P}'' = \mathcal{P} \setminus \mathcal{P}'$.

It is clear that the family of triples $\{\Pi_1, \Pi_2, \Pi_3\}$, as Π ranges over $\mathcal{P}'$, is a compatible family of 3-partitions of Z, whence

$$|\mathcal{P}'| \le |Z| - 2 = |Y| - 3,$$

by the induction hypothesis.

Finally we show that $|\mathcal{P}''| \le 1$. Suppose, on the contrary, that $\mathcal{P}''$ contains two distinct 3-partitions of Y, which we write as $\{\{y\}, A_2, A_3\}$ and $\{\{y\}, B_2, B_3\}$. By a suitable re-ordering of the elements, we may assume that $A_2 \supseteq \{y\} \cup B_2$, which is a contradiction since $y \notin A_2$.

In summary,

$$|\mathcal{P}| \le |\mathcal{P}'| + |\mathcal{P}''| \le |Y| - 2$$

as required. □

Let K be a connected open cluster of $B(n)$. If x is a trifurcation belonging to $K \cap B(n-1)$, the removal of x induces a partition of $K \cap \partial B(n)$ into three sets, namely the sets joined by open paths to x via each of the three open edges incident to x. Therefore, x corresponds to a 3-partition $\Pi(x) = \{P_1, P_2, P_3\}$ of $K \cap \partial B(n)$ with the properties that:

(a) P_i is non-empty, for $i = 1, 2, 3$;

(b) P_i is a subset of a connected open subgraph of $B(n) \setminus \{x\}$;

(c) $P_i \nleftrightarrow P_j$ in $B(n) \setminus \{x\}$, if $i \neq j$.

Furthermore, if x and x' are distinct trifurcations of $K \cap B(n-1)$, then $\Pi(x)$ and $\Pi(x')$ are distinct and compatible; see Figure 8.2.

It follows by Lemma 8.5 that the number $\tau(K)$ of trifurcations in $K \cap B(n-1)$ satisfies

$$\tau(K) \le |K \cap \partial B(n)| - 2. \tag{8.6}$$

We sum this inequality over all connected clusters K of $B(n)$, to obtain that

$$\sum_{x \in B(n-1)} I_{T_x} \le |\partial B(n)|.$$

Take expectations and use (8.4), to find that

$$|B(n-1)| P_p(T_0) \le |\partial B(n)|, \tag{8.7}$$

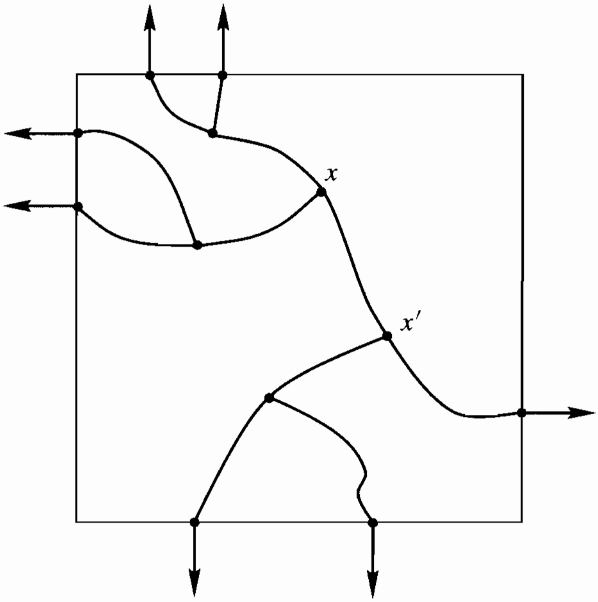


Figure 8.2. Two trifurcations x and x' belonging to a cluster K of $B(n)$. They induce compatible partitions of $K \cap \partial B(n)$.

which is impossible for large n since the left side grows as n^d and the right side as n^{d-1}. This contradiction completes the proof. □

8.3 Continuity of the Percolation Probability

Certainly the percolation probability is continuous when $p < p_c$, since $\theta(p) = 0$ in this case. It is a consequence of the uniqueness of the infinite open cluster that θ is continuous when $p > p_c$ also.

(8.8) Theorem. Continuity of θ in the supercritical phase. *The percolation probability $\theta(p)$ is a continuous function of p on the interval $(p_c, 1]$.*

It is an open question whether or not θ is continuous at p_c. It is not difficult to show that θ is continuous from the right on the whole interval $[0, 1]$, and thus

$$\theta(p_c) = \lim_{p \downarrow p_c} \theta(p).$$

On the other hand, $\theta(p) = 0$ for $p < p_c$, giving that θ is continuous at p_c if and only if $\theta(p_c) = 0$, which is to say that all open clusters are finite almost surely when $p = p_c$. It is believed that $\theta(p_c) = 0$ for all $d \geq 2$, but this is known to hold only when $d = 2$ or $d \geq 19$ (see the notes for this section). In this direction, it is known that the percolation probability is continuous *at* the critical point for percolation on a half-space; see Section 7.3. In Section 8.7 we shall see that θ is infinitely differentiable when $p > p_c$; is is believed that θ is real analytic on $(p_c, 1]$.

We prove the above theorem in two parts.

(8.9) Lemma. *The percolation probability function $\theta(p)$ is continuous from the right on the interval* $[0, 1]$.

(8.10) Lemma. *The percolation probability function $\theta(p)$ is continuous from the left on the interval* $(p_c, 1]$.

Proof of Lemma (8.9). We follow Russo (1978), using an argument previewed in the proof of Theorem (5.8). Clearly,

$$\theta(p) = \lim_{n\to\infty} P_p\big(0 \leftrightarrow \partial B(n)\big).$$

However, the function $P_p(0 \leftrightarrow \partial B(n))$ varies continuously with p, since the event $\{0 \leftrightarrow \partial B(n)\}$ depends only on the finite set of edges in $B(n)$; moreover,

$$P_p\big(0 \leftrightarrow \partial B(n)\big) \geq P_p\big(0 \leftrightarrow \partial B(n+1)\big) \qquad \text{for all } n.$$

Thus θ is the decreasing limit of continuous functions, and therefore θ is upper semi-continuous. On the other hand, θ is monotonic non-decreasing, and thus θ is continuous from the right. □

Before moving to the proof of Lemma (8.10), we note that the above argument may be used to establish the continuity in p of the connectivity function $\tau_p(x, y)$. Let

$$\tau_p\big(x, y; B(n)\big) = P_p\big(x \leftrightarrow y \text{ in } B(n)\big)$$

be the probability that x and y are joined by an open path of $B(n)$. It is clear that

$$\tau_p\big(x, y; B(n)\big) \leq \tau_p\big(x, y; B(n+1)\big),$$

so that

$$\tau_p\big(x, y; B(n)\big) \uparrow \tau_p(x, y) \qquad \text{as } n \to \infty.$$

Now $\tau_p(x, y; B(n))$ is a continuous function of p, and therefore $\tau_p(x, y)$ is lower semi-continuous.

For the upper bound on $\tau_p(x, y)$, we use

(8.11)
$$\tau_p^+\big(x, y; B(n)\big) = P_p\Big(\big\{x \leftrightarrow y \text{ in } B(n)\big\} \cup \big\{x \leftrightarrow \partial B(n) \text{ and } y \leftrightarrow \partial B(n)\big\}\Big),$$

the probability that either x and y are in the same open cluster of $B(n)$, or both are joined by open paths to the surface $\partial B(n)$ of $B(n)$. It is easy to see that

$$\tau_p^+\big(x, y; B(n)\big) \ge \tau_p^+\big(x, y; B(n+1)\big)$$

and

$$\tau_p^+\big(x, y; B(n)\big) \to \tau_p^+(x, y) \qquad \text{as } n \to \infty, \tag{8.12}$$

where

$$\tau_p^+(x, y) = P_p\,(\text{either } x \leftrightarrow y, \text{ or } x \text{ and } y \text{ belong to different infinite open clusters}). \tag{8.13}$$

There is probability 0 of two or more infinite open clusters, so that $\tau_p^+(x, y) = \tau_p(x, y)$, giving that $\tau_p(x, y)$ is the decreasing limit of $\tau_p^+(x, y; B(n))$ as $n \to \infty$. The latter quantities are continuous functions of p, and thus $\tau_p(x, y)$ is upper semi-continuous. Hence $\tau_p(x, y)$ is a continuous function of p. The same argument is valid for the m-point connectivity function $\tau_p(x_1, x_2, \dots, x_m)$.

Proof of Lemma (8.10). We follow van den Berg and Keane (1984), and shall use the uniqueness theorem for the infinite open cluster. Let $(X(e) : e \in \mathbb{E}^d)$ be a collection of independent random variables indexed by the set of edges of $\mathbb{L}^d$, each having the uniform distribution on $[0, 1]$. For $0 \le p \le 1$, we define

$$\eta_p(e) = \begin{cases} 1 & \text{if } X(e) < p, \\ 0 & \text{otherwise.} \end{cases}$$

We say that an edge e is *p-open* if $\eta_p(e) = 1$ and *p-closed* otherwise. We write C_p for the p-open cluster of $\mathbb{L}^d$ containing the origin, and note that $C_\pi \subseteq C_p$ if $\pi \le p$. Now,

$$\theta(p) = P(|C_p| = \infty) \tag{8.14}$$

and

$$\lim_{\pi \uparrow p} \theta(\pi) = \lim_{\pi \uparrow p} P(|C_\pi| = \infty) = P\big(|C_\pi| = \infty \text{ for some } \pi < p\big), \tag{8.15}$$

since the event $\{|C_\pi| = \infty\}$ is increasing in π. Bearing in mind the monotonicity of θ, we are required to prove that

$$P\big(|C_p| = \infty,\ |C_\pi| < \infty \text{ for all } \pi < p\big) = 0 \qquad \text{for } p > p_c. \tag{8.16}$$

Let $p > p_c$, and suppose that $|C_p| = \infty$. If $p_c < \alpha < p$, there exists (almost surely) an infinite α-open cluster I_α, and furthermore I_α is almost surely a subgraph of C_p, since otherwise there would exist at least two infinite p-open clusters. It follows that there exists a p-open path l joining the origin to some vertex of I_α. Such a path l has finite length and each edge e in l satisfies $X(e) < p$; therefore $\mu = \max\{X(e) : e \in l\}$ satisfies $\mu < p$. If π satisfies $\pi \ge \alpha$ and $\mu < \pi < p$, there exists a π-open path joining the origin to some vertex of I_α, so that $|C_\pi| = \infty$, implying that (8.16) holds as required. □

8.4 The Radius of a Finite Open Cluster

When $p < p_c$, the probability of an open path from the origin to a vertex on the surface $\partial B(n)$ of the box $B(n)$ decays exponentially as $n \to \infty$. This is of course false when $p > p_c$, since this probability is then at least as large as the percolation probability $\theta(p)$. We therefore turn our attention to the probability that there exists an open path from the origin to $\partial B(n)$ *not lying in the infinite open cluster*. We may expect this probability to decay exponentially as $n \to \infty$, in that there exists $\sigma(p)$ such that

$$(8.17) \qquad P_p\big(0 \leftrightarrow \partial B(n),\ |C| < \infty\big) \approx e^{-n\sigma(p)} \qquad \text{as } n \to \infty.$$

(Remember that $a_n \approx b_n$ means $\log a_n / \log b_n \to 1$.) We may expect further that $\sigma(p) > 0$ when $p > p_c$. The next theorem establishes the existence of the limit required for (8.17).

(8.18) Theorem. Asymptotic tail behaviour of the radius of a finite open cluster. *Suppose that* $0 < p < 1$. *The limit*

$$(8.19) \qquad \sigma(p) = \lim_{n\to\infty} \left\{ -\frac{1}{n} \log P_p\big(0 \leftrightarrow \partial B(n),\ |C| < \infty\big) \right\}$$

exists and satisfies $\sigma(p) < \infty$. *Furthermore, there exists a constant* $A(p,d)$ *which is finite for* $d \ge 2$, $0 < p < 1$, *such that*

$$(8.20) \qquad P_p\big(0 \leftrightarrow \partial B(n),\ |C| < \infty\big) \le A(p,d)n^d e^{-n\sigma(p)} \qquad \textit{for all } n.$$

This theorem is proved using subadditivity in a way not dissimilar to the proof of Lemma (6.102), although there are certain extra complications owing to the condition that the open cluster at the origin be finite. The main question is, of course, to determine those values of p for which $\sigma(p)$ is strictly positive, since only then does the tail of the radius distribution decay exponentially.

(8.21) Theorem. *We have that* $\sigma(p) > 0$ *if* $p > p_c$. *Furthermore,* $\sigma(p)$ *is uniformly bounded away from* 0 *on any closed sub-interval of* $(p_c, 1)$.

We shall prove Theorem (8.21) under the assumption that $d \ge 3$, returning in Theorem (11.24) to the case $d = 2$.

One important technique in the proof of Theorem (8.21) is a method for constructing the open cluster C at the origin in a step by step manner, and we do this as follows. First we order the edges of $\mathbb{L}^d$ in some arbitrary but deterministic way, and we write e_i for the ith edge in this ordering. We shall construct an increasing sequence $C_1, C_2, \dots$ of (random) subgraphs of $\mathbb{L}^d$. We define $C_1 = \{0\}$, the graph containing the origin only. Having found C_m, we let C_{m+1} be the graph obtained by adding to C_m the earliest open edge which lies in the edge boundary of C_m, if such an edge exists. That is to say, we define

$$(8.22) \qquad C_{m+1} = C_m \cup \{e_j\},$$

where

$$j = \min\{i : e_i \notin C_m,\ e_i \text{ is open and incident to a vertex of } C_m\}. \tag{8.23}$$

Thus, we add to C_m the edge e_j together with any endvertex of e_j which does not already belong to C_m. A single edge is added at each stage, until the stage when no more suitable open edges can be found. This occurs when we have exhausted the entire open cluster C at the origin. If $C_m = C$ for some C then we define $C_l = C$ for $l \geq m$, so that

$$C = \lim_{m\to\infty} C_m. \tag{8.24}$$

The main property of this construction is that the C_m have a certain property similar to that of stopping times for a Markov process: for any m and any connected subgraph Σ of $\mathbb{L}^d$ containing the origin, the event $\{C_m = \Sigma\}$ depends only on the states of edges in Σ and its edge boundary, and not on the state of any edge having both endvertices outside Σ. We note that, if x is any vertex of the open cluster C at the origin, there exists a random integer $m = m(x)$ such that x lies in C_m, regardless of whether C is finite or infinite.

Proof of Theorem (8.18). We follow Chayes, Chayes, Grimmett, Kesten, and Schonmann (1989). First we shall prove the existence of the limit in (8.19), and we begin with some notation. For $i = 1, 2, \dots, d$, we define

$$R_i = \max\{x_i : x \in C\}, \qquad L_i = \min\{x_i : x \in C\}, \tag{8.25}$$

the maximum and minimum values of the ith coordinates of vertices in the open cluster C at the origin. We write $D_i = R_i - L_i$ for the *width* of C in the ith coordinate direction, and we define the *diameter* of C by

$$\text{diam}(C) = \max\{D_i : 1 \leq i \leq d\}. \tag{8.26}$$

(8.27) Lemma. *For $m, n \geq 0$,*

(8.28)
$$P_p\big(\text{diam}(C) = m + n + 2\big) \geq \frac{p^2(1-p)^{2d-2}}{d^2(2n+1)^d} P_p\big(\text{diam}(C) = m\big) P_p\big(\text{diam}(C) = n\big).$$

Proof. The idea of the proof is as follows. If we are provided with two finite open clusters whose widths in the first coordinate direction are m and n, then we may position one of them to the right of the other in order to obtain a finite open cluster with width $m + n + 2$, the extra 2 arising out of the geometry of the construction, illustrated in Figure 8.3.

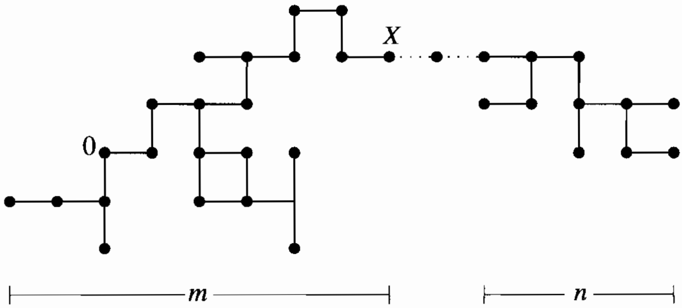


Figure 8.3. Two open clusters having widths m and n in the first coordinate direction may be used to build an open cluster having width $m+n+2$.

First we note that, if $\operatorname{diam}(C) = k$, we do not lose a great deal by assuming that the width of C is maximal in the first coordinate direction. Certainly

$$(8.29) \qquad P_p\big(\operatorname{diam}(C) = k\big) \geq P_p\big(D_1 = \operatorname{diam}(C) = k\big),$$

and we claim also that

$$(8.30) \qquad P_p\big(D_1 = \operatorname{diam}(C) = k\big) \geq \frac{1}{d} P_p\big(\operatorname{diam}(C) = k\big).$$

It is the case that $\operatorname{diam}(C) = D_i$ for some $i = 1, 2, \ldots, d$, and therefore, if (8.30) were false, it would follow by symmetry that

$$(8.31) \quad P_p\big(\operatorname{diam}(C) = k\big) \leq \sum_{i=1}^{d} P_p\big(D_i = \operatorname{diam}(C) = k\big) < P_p\big(\operatorname{diam}(C) = k\big),$$

a contradiction.

Suppose that $D_1 = \operatorname{diam}(C) = k$. Then $C \subseteq B(k)$, the box with side-length $2k$ and centre at the origin. The cluster C has at least one *leftmost* vertex (that is, a vertex v with $v_1 \leq w_1$ for all vertices w in C), and it follows that there exists a (non-random) vertex x of $B(k)$ with the property that

$$(8.32) \qquad P_p\big(x \text{ is a leftmost vertex of } C,\ D_1 = \operatorname{diam}(C) = k\big) \geq \frac{1}{|B(k)|} P_p\big(D_1 = \operatorname{diam}(C) = k\big).$$

To see this we argue as before. If (8.32) were false for all $x \in B(k)$ then

$$P_p\big(D_1 = \operatorname{diam}(C) = k\big) \leq \sum_{x \in B(k)} P_p\,(x \text{ is a leftmost vertex of } C,\ D_1 = \operatorname{diam}(C) = k) < P_p\big(D_1 = \operatorname{diam}(C) = k\big),$$

a contradiction. If $D_1 = \operatorname{diam}(C) = k$ and x is a leftmost vertex of C, then x is a leftmost vertex of an open cluster with diameter k whose width is maximal in the first coordinate direction. The probability of this event is constant for all vertices x, and thus we may take x to be the origin, obtaining that

$$(8.33) \qquad P_p\big(D_1 = \operatorname{diam}(C) = k,\ L_1 = 0,\ R_1 = k\big) \geq \frac{1}{|B(k)|} P_p\big(D_1 = \operatorname{diam}(C) = k\big).$$

When combined with (8.30), this implies that

$$(8.34) \qquad P_p\big(\operatorname{diam}(C) = k,\ L_1 = 0,\ R_1 = k\big) \geq \frac{1}{d|B(k)|} P_p\big(\operatorname{diam}(C) = k\big).$$

We perform next the required geometrical construction, illustrated in Figure 8.3. Let m and n be non-negative integers, and suppose that the open cluster C at the origin is such that $D_1 = \operatorname{diam}(C) = m$; that is, we assume that C has diameter m and the width of C is maximal in the first coordinate direction. We call a vertex x of C *rightmost* if $y_1 \leq x_1$ for all $y \in C$, and we denote by X a rightmost vertex of C; if there exists more than one such vertex, we pick that with greatest second coordinate, and if there is more than one of these then we pick the one with greatest third coordinate, and so on. Having found X, we consider the event that the vertex $X + (2, 0, 0, \dots, 0)$ is a leftmost vertex of the open cluster containing it, and that this open cluster has diameter n and its width is maximal in the first coordinate direction. By (8.34), the probability of this event, conditional on X, is at least $(d|B(n)|)^{-1} P_p(\operatorname{diam}(C) = n)$; we have no extra information about the states of edges for which both endvertices z satisfy $z_1 > x_1$. We now change the states of the edges incident to $X + (1, 0, 0, \dots, 0)$ so that the edges leading to X and $X + (2, 0, 0, \dots, 0)$ are open but all other incident edges are closed. The outcome of this construction is a finite open cluster containing the origin and having diameter $m + n + 2$; furthermore, any such outcome arises in a unique way from suitable clusters with diameters m and n. Thus, the probability that C has diameter $m + n + 2$ is at least

$$(8.35) \qquad P_p\big(D_1 = \operatorname{diam}(C) = m\big) \frac{1}{d|B(n)|} P_p\big(\operatorname{diam}(C) = n\big) p^2(1-p)^{2d-2},$$

where the term $p^2(1-p)^{2d-2}$ is the worst penalty payable for the changes in the states of edges incident to $X + (1, 0, 0, \dots, 0)$. Thus

$$\begin{aligned} P_p\big(\operatorname{diam}(C) = m + n + 2\big) &\geq \frac{p^2(1-p)^{2d-2}}{d|B(n)|} P_p\big(D_1 = \operatorname{diam}(C) = m\big) P_p\big(\operatorname{diam}(C) = n\big) \\ &\geq \frac{p^2(1-p)^{2d-2}}{d^2|B(n)|} P_p\big(\operatorname{diam}(C) = m\big) P_p\big(\operatorname{diam}(C) = n\big) \end{aligned}$$

by (8.30). Of course, $|B(n)| = (2n+1)^d$. □

Let us now concentrate on inequality (8.28). Taking logarithms, we find that the sequence

$$\delta(k) = -\log P_p\big(\text{diam}(C) = k\big)$$

satisfies

$$\delta(m+n+2) \le \delta(m) + \delta(n) + \log\left\{\frac{d^2(2n+1)^d}{p^2(1-p)^{2d-2}}\right\}. \tag{8.36}$$

This differs slightly from the usual form of the generalized subadditive inequality in that the argument on the left side is $m+n+2$ instead of $m+n$, and the error term is a function n rather than of $m+n$. These minor aberrations are easily taken into account in the proof of the subadditive limit theorem. As noted at the end of Appendix II, inequality (8.36) implies the existence of the limit

$$\sigma(p) = \lim_{n\to\infty}\left\{\frac{1}{n}\,\delta(n)\right\} = \lim_{n\to\infty}\left\{-\frac{1}{n}\log P_p\big(\text{diam}(C) = n\big)\right\} \tag{8.37}$$

together with the inequality

$$\delta(k) \ge (k+2)\sigma(p) - \log\left\{\frac{d^2(2k+1)^d}{p^2(1-p)^{2d-2}}\right\} \qquad \text{for all } k,$$

which we rewrite as

$$P_p\big(\text{diam}(C) = k\big) \le \frac{d^2}{p^2(1-p)^{2d-2}}\,(2k+1)^d e^{-(k+2)\sigma(p)}. \tag{8.38}$$

In particular, $\sigma(p) < \infty$ if $0 < p < 1$.

Having proved the exponential behaviour of the tail of the diameter distribution, it is easy to show the same for the radius. It is clear that

$$P_p\big(0 \leftrightarrow \partial B(n),\ |C| < \infty\big) \le P_p\big(n \le \text{diam}(C) < \infty\big), \tag{8.39}$$

since the diameter of C is at least as large as its radius. On the other hand,

$$P_p\big(0 \leftrightarrow \partial B(n),\ |C| < \infty\big) \ge P_p\big(\text{diam}(C) = n,\ L_1 = 0,\ R_1 = n\big),$$

since, if the origin is a leftmost vertex of a finite cluster with width n in the first coordinate direction, there exists an open path joining the origin to $\partial B(n)$. Hence

$$P_p\big(0 \leftrightarrow \partial B(n),\ |C| < \infty\big) \ge \frac{1}{d(2n+1)^d}\,P_p\big(\text{diam}(C) = n\big) \tag{8.40}$$

by (8.34). It follows from (8.39) and (8.40), in conjunction with (8.37), that

$$(8.41) \qquad -\frac{1}{n} \log P_p\big(0 \leftrightarrow \partial B(n),\ |C| < \infty\big) \to \sigma(p) \qquad \text{as } n \to \infty.$$

The only case requiring special attention is when $\sigma(p) = 0$, in which instance (8.40) alone implies (8.41).

Finally, we show that there exists a constant $A(p, d)$ such that (8.20) holds. Suppose that $\sigma(p) > 0$. From (8.38) and (8.39),
(8.42)

$$P_p\big(0 \leftrightarrow \partial B(n),\ |C| < \infty\big) \le \frac{d^2}{p^2(1-p)^{2d-2}} \sum_{m=n}^{\infty} (2m+1)^d e^{-(m+2)\sigma(p)},$$

of which (8.20) is an immediate consequence with a suitable choice for $A(p, d)$. If $\sigma(p) = 0$ then (8.20) is satisfied with $A(p, d) = 1$. □

Proof of Theorem (8.21). We turn to the proof that $\sigma(p) > 0$ when $p > p_c$ and $d \ge 3$, and we shall follow Chayes, Chayes, and Newman (1987). Let $H(n)$ denote the hyperplane $\{x \in \mathbb{Z}^d : x_1 = n\}$ of $\mathbb{Z}^d$, and let G_n be the event that the origin belongs to a finite cluster which intersects $H(n)$. The box $B(n)$ has $2d$ faces, and so

$$(8.43) \qquad P_p(G_n) \le P_p\big(0 \leftrightarrow \partial B(n),\ |C| < \infty\big) \le 2d P_p(G_n),$$

as in (6.22). We shall prove that, if $p > p_c$, there exists $\gamma(p) > 0$ such that

$$(8.44) \qquad P_p(G_n) \le e^{-n\gamma(p)} \qquad \text{for all } n;$$

it follows from the second inequality in (8.43) that $\sigma(p) \ge \gamma(p)$, and it is an immediate consequence that $\sigma(p) > 0$ when $p > p_c$.

The idea of the proof is as follows. Let

$$R_k = \big\{x \in \mathbb{Z}^d : 0 \le x_1 < k\big\}.$$

Since $p > p_c$, we may by Theorem (7.2) choose an integer k such that $p > p_c(R_k)$. Suppose that G_{nk} occurs for some $n \ge 1$. In this case, each of the regions

$$(8.45) \qquad R_k(i) = \big\{x \in \mathbb{Z}^d : (i-1)k \le x_1 < ik\big\},$$

for $1 \le i \le n$, is traversed by an open path from the origin, and this open path does not intersect any infinite open cluster contained within such regions; see Figure 8.4. Since $p > p_c(R_k)$, each region contains a.s. an infinite open cluster.

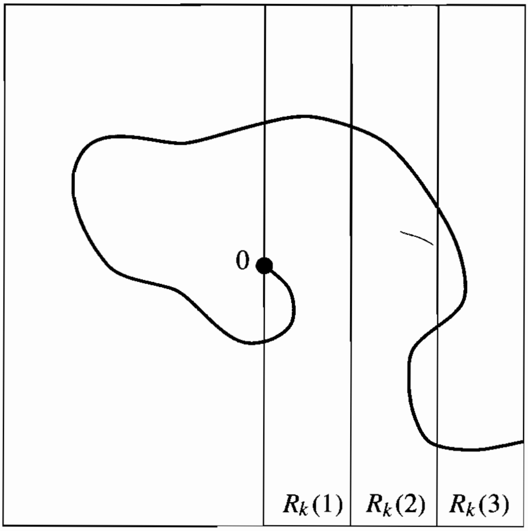


Figure 8.4. All paths from the origin to $H(3k)$ traverse the regions $R_k(1)$, $R_k(2)$, and $R_k(3)$.

and the probability of avoiding it is strictly less than 1. There are n such regions, giving that the probability of G_{nk} is no larger than α^n for some $\alpha = \alpha(p, k) < 1$.

We have to do a certain amount of work to make this argument rigorous. Let k be a positive integer. First, we construct the open cluster C at the origin as the limit of the sequence $(C_m : m \geq 1)$ in the manner described just after the statement of the theorem. Next we construct a sequence $v_1, v_2, \dots$ of vertices in the following manner. We set $v_1 = 0$, the origin. We let m_i be the smallest value of m for which C_m contains a vertex of the region $R_k(i)$, and we denote this vertex by v_i. Such a v_i exists if and only if some vertex in $R_k(i)$ lies in the open cluster at the origin. In this way, we obtain a sequence $v_1, v_2, \dots, v_T$ of vertices where

$$T = \sup\{i : C \cap R_k(i) \neq \varnothing\}.$$

Let $\theta(p, v; R)$ denote the probability that the vertex v belongs to an infinite open cluster of the region R. We choose k such that

$$p > p_c(R_k) = \sup\{p : \theta(p, 0; R_k) = 0\}.$$

Note that $\theta(p, 0; R_k) > 0$. Let n be a positive integer, and write $n = kr + s$ where $0 \leq s < k$. If the event G_n occurs then $T \geq r$ and v_i does not lie in an infinite open cluster of $R_k(i)$ for $1 \leq i \leq r$. Thus

$$(8.46)\quad \begin{aligned} P_p(G_n) &\leq P_p\big(T \geq r,\ v_i \text{ is not in an infinite open cluster of } R_k(i), \\ &\qquad \text{for } 1 \leq i \leq r\big) \\ &= P_p(A_r) \\ &= P_p(A_r \mid A_{r-1}) P_p(A_{r-1}), \end{aligned}$$

where

$$A_j = \{T \geq j\} \cap \left\{ \bigcap_{1 \leq i \leq j} \{v_i \text{ does not lie in an infinite cluster of } R_k(i)\} \right\}.$$

By the method of construction of the C_m, the events A_{j-1} and $\{T \geq j\}$ are independent of the states of all edges joining vertices in $R_k(j)$, whereas, conditional on v_j, the event that v_j lies in the infinite open cluster of $R_k(j)$ depends *only* on edges joining vertices in $R_k(j)$. Therefore,

$$\begin{aligned}
&P_p(A_j \mid A_{j-1}) \\
&\quad \leq \sum_v P_p\big(v \text{ not in an infinite open cluster of } R_k(j) \,\big|\, v_j = v,\ T \geq j,\ A_{j-1}\big) \\
&\qquad\qquad\qquad \times P_p\big(v_j = v \,\big|\, T \geq j,\ A_{j-1}\big) P_p\big(T \geq j \,\big|\, A_{j-1}\big) \\
&\quad \leq \sum_v \big(1 - \theta(p, v; R_k(j))\big) P_p\big(v_j = v \,\big|\, T \geq j,\ A_{j-1}\big) \\
&\quad = 1 - \theta(p, 0; R_k),
\end{aligned}$$

where the sum is over all vertices v in the hyperplane $H((j-1)k)$. This is valid for all $j \geq 2$, giving from (8.46) that

$$\begin{aligned}
P_p(G_n) &\leq P_p(A_r \mid A_{r-1}) P_p(A_{r-1} \mid A_{r-2}) \dots P_p(A_2 \mid A_1) P_p(A_1) \\
&\leq \big(1 - \theta(p, 0; R_k)\big)^r.
\end{aligned}$$

However, $\theta(p, 0; R_k) > 0$, and thus

(8.47) $$P_p(G_n) \leq \exp\big(-r\gamma_k(p)\big)$$

where

$$\gamma_k(p) = -\log\big(1 - \theta(p, 0; R_k)\big) > 0.$$

Now $n \leq (r+1)k$, so that $r \geq (n/k) - 1$, giving that

(8.48) $$P_p(G_n) \leq \exp\big(-n\{\gamma_k(p)/k\} + \gamma_k(p)\big)$$

for all $n \geq 0$. This proves that $\sigma(p) > 0$.

Finally, let $p_c < \alpha < \beta < 1$, and choose k such that $p_c(R_k) < \alpha$. We have by (8.48) that

$$P_p(G_n) \leq \exp\big(-n\{\gamma_k(\alpha)/k\} + \gamma_k(\beta)\big).$$

Therefore, $\sigma(p)$ is uniformly bounded away from 0 on the interval $[\alpha, \beta]$. □

8.5 Truncated Connectivity Functions and Correlation Length

The two-point connectivity function τ_p is defined by $\tau_p(x, y) = P_p(x \leftrightarrow y)$. When $p < p_c$, $\tau_p(x, y)$ decays roughly as an exponential function of the distance $\delta(x, y)$ from x to y. When $p > p_c$, $\tau_p(x, y)$ does not decay at all, since

$$\begin{aligned}\tau_p(x, y) &\geq P_p\big(x \text{ and } y \text{ belong to the unique infinite open cluster}\big)\\ &= P_p\big(|C(x)| = \infty, |C(y)| = \infty\big)\\ &\geq P_p(|C(x)| = \infty)P_p(|C(y)| = \infty) = \theta(p)^2,\end{aligned}$$

where we have used the FKG inequality. It is an open question to determine the behaviour of $\tau_p(x, y)$ when $p = p_c$.

Of more interest than τ_p when $p > p_c$ is the *truncated connectivity function* τ_p^f defined by

$$\tau_p^f(x, y) = P_p\big(x \text{ and } y \text{ belong to the same finite open cluster}\big). \tag{8.49}$$

Note that

$$\tau_p^f(x, y) = \tau_p(x, y) - P_p\big(x \text{ and } y \text{ belong to the same infinite open cluster}\big),$$

so that

$$\tau_p(x, y) = \tau_p^f(x, y) \qquad \text{if } \theta(p) = 0. \tag{8.50}$$

The superscript 'f' refers to the requirement that x and y be in *finite* open clusters, just as $\chi^f(p)$ denotes the mean size of the *finite* open cluster at the origin.

In the next theorem, we prove that $\tau_p^f(x, y)$ decays in the manner of an exponential function of the distance between x and y. It is an immediate consequence of Theorem (8.18) that there exists a constant $A(p, d)$, finite for $d \geq 2$ and $0 < p < 1$, such that

$$\tau_p^f(0, x) \leq A(p, d)\|x\|^d e^{-\|x\|\sigma(p)} \qquad \text{for all } x, \tag{8.51}$$

where $\sigma(p)$ is given in (8.19) and $\|x\| = \max\{|x_i| : 1 \leq i \leq d\}$; we recall that $\sigma(p) > 0$ if $p > p_c$. There is a formula for $\sigma(p)$ in two dimensions: we shall see in Theorem (11.24) that, if $d = 2$,

$$-\frac{1}{n}\log \tau_p^f(0, e_n) \to \frac{2}{\xi(1-p)} \qquad \text{as } n \to \infty, \tag{8.52}$$

where $e_n = (n, 0)$ and $\xi(1-p)$ is the correlation length of the subcritical percolation process with edge-probability $1 - p$. No such simple formula is likely to be valid when $d \geq 3$.

In studying the truncated connectivity function, we shall concentrate on the event that the origin and the vertex $e_n = (n, 0, 0, \dots, 0)$ belong to the same finite open cluster. Our principal result states that the probability of this event has the same asymptotic behaviour as the probability that the radius of the open cluster at the origin is finite but at least n.

(8.53) Theorem. Asymptotic behaviour of truncated connectivity function. *Suppose that* $0 < p < 1$, *and let* $\sigma(p)$ *and* $A(p, d)$ *be as given in Theorem* (8.18). *Then*

(8.54) $$\lim_{n\to\infty}\left\{-\frac{1}{n}\log\tau_p^{\mathrm{f}}(0, e_n)\right\} = \sigma(p)$$

and

(8.55) $$\tau_p^{\mathrm{f}}(0, e_n) \le A(p, d)n^d e^{-n\sigma(p)} \qquad \textit{for all } n.$$

In the subcritical phase ($p < p_c$), the correlation length $\xi(p)$ was given by

$$\begin{aligned}\xi(p)^{-1} &= \lim_{n\to\infty}\left\{-\frac{1}{n}\log P_p\big(0 \leftrightarrow \partial B(n)\big)\right\}\\ &= \lim_{n\to\infty}\left\{-\frac{1}{n}\log\tau_p(0, e_n)\right\}.\end{aligned}$$

We define the *correlation length* $\xi(p)$ analogously in the supercritical phase,

(8.56) $$\begin{aligned}\xi(p)^{-1} &= \lim_{n\to\infty}\left\{-\frac{1}{n}\log P_p\big(0 \leftrightarrow \partial B(n),\ |C| < \infty\big)\right\}\\ &= \lim_{n\to\infty}\left\{-\frac{1}{n}\log\tau_p^{\mathrm{f}}(0, e_n)\right\},\end{aligned}$$

by reference to the corresponding events involving *finite* open clusters. This seems to be a sensible definition, and we shall return to the notion of correlation length in Chapter 9.

Proof of Theorem (8.53). We follow Chayes, Chayes, Grimmett, Kesten, and Schonmann (1989). The upper bound on $\tau_p^{\mathrm{f}}(0, e_n)$ is easy. Just note that

(8.57) $$\tau_p^{\mathrm{f}}(0, e_n) \le P_p\big(0 \leftrightarrow \partial B(n),\ |C| < \infty\big),$$

and use Theorem (8.18). Inequality (8.51) is derived similarly.

Slightly more work is required for the lower bound, and we follow the line of argument in the proof of Theorem (6.44). Let m be a positive integer, and suppose that the open cluster C at the origin has the following properties:

(a) $\mathrm{diam}(C) = D_1$, so that the width of C is maximal in the first coordinate direction;

(b) $\mathrm{diam}(C) = m$;

(c) $L_1 = 0$ and $R_1 = m$, so that the origin is a leftmost vertex of C in the sense that $x_1 \ge 0$ for all $x \in C$;

where we use the notation of the last section. Under these conditions, there exists a vertex in the hyperplane $H(m) = \{x \in \mathbb{Z}^d : x_1 = m\}$ which is in the same finite open cluster as the origin, and furthermore this vertex is a rightmost vertex of C. In a similar way to (8.30) and (8.32), there exists a (non-random) vertex x in $H(m) \cap \partial B(m)$ for which

$$(8.58) \quad P_p\big(x \in C, \operatorname{diam}(C) = m,\ L_1 = 0,\ R_1 = m\big) \geq \frac{1}{|\partial B(m)|} P_p\big(\operatorname{diam}(C) = m,\ L_1 = 0,\ R_1 = m\big),$$

and we choose x accordingly. We now put two such events together in such a way as to force the origin and the vertex e_{2m+2} to be in the same finite open cluster. Let us consider the event that e_{2m+2} and the vertex $x + (2, 0, 0, \dots, 0)$ are (respectively) rightmost and leftmost vertices of the same finite open cluster having diameter m; by symmetry, this event has the same probability as the left side of equation (8.58), and furthermore it is independent of the event that the origin and x are leftmost and rightmost vertices of the same finite open cluster having diameter m. If we change the states of the edges incident to $x + (1, 0, 0, \dots, 0)$ so that the two edges leading to x and to $x + (2, 0, 0, \dots, 0)$ are open but all others are closed, then we force the origin and e_{2m+2} to be in the same finite open cluster. We have proved that

(8.59)

$$\begin{aligned}
\tau_p^{\mathrm{f}}(0, e_{2m+2}) &\geq p^2(1-p)^{2d-2} P_p\big(x \in C, \operatorname{diam}(C) = m,\ L_1 = 0,\ R_1 = m\big)^2 \\
&\geq \frac{p^2(1-p)^{2d-2}}{|\partial B(m)|^2} P_p\big(\operatorname{diam}(C) = m,\ L_1 = 0,\ R_1 = m\big)^2 \\
&\geq \frac{p^2(1-p)^{2d-2}}{d^2|\partial B(m)|^2|B(m)|^2} P_p\big(\operatorname{diam}(C) = m\big)^2
\end{aligned}$$

by (8.58) and (8.34). In a similar way, we have that

$$(8.60) \quad \tau_p^{\mathrm{f}}(0, e_{2m+3}) \geq \frac{p^3(1-p)^{4d-4}}{d^2|\partial B(m)|^2|B(m)|^2} P_p\big(\operatorname{diam}(C) = m\big)^2.$$

With the aid of (8.19) and (8.37), inequalities (8.57), (8.59), and (8.60) imply jointly that (8.54) holds. □

8.6 Sub-Exponential Decay of the Cluster Size Distribution

We turn our attention next to the number of vertices in the open cluster at the origin, when this cluster is finite. Whereas $P_p(|C| = n)$ decays exponentially in n when $p < p_c$, it has a quite different behaviour when $p > p_c$. It is fairly easy to see why the following result should be valid, although some technical complications arise in its proof.

(8.61) Theorem. Sub-exponential decay of the cluster size distribution. *Suppose that $p_c < p < 1$. There exists $\gamma(p) < \infty$ such that*

$$P_p(|C| = n) \geq \exp\big(-\gamma(p)n^{(d-1)/d}\big) \qquad \textit{for all } n. \tag{8.62}$$

We recall that $n^{(d-1)/d}$ is the order of the surface area of a ball of $\mathbb{R}^d$ with volume n. We say that the cluster size distribution decays 'sub-exponentially' since it decays to 0 at a rate which is slower than exponential decay; some refer to this decay as 'stretched exponential'.

It is believed that this theorem describes the correct order of magnitude of $P_p(|C| = n)$, and we therefore make the conjecture that the limit

$$\delta(p) = \lim_{n\to\infty} \big\{-n^{-(d-1)/d} \log P_p(|C| = n)\big\} \tag{8.63}$$

exists and satisfies $0 < \delta(p) < \infty$ when $p_c < p < 1$. Assuming the existence of the limit $\delta(p)$, one would like to understand how it behaves as $p \downarrow p_c$; at what rate does $\delta(p) \to 0$ as $p \downarrow p_c$? The existence of the limit in (8.63) has been proved when $d = 2, 3$; see the notes for this section.

We turn next to upper bounds for $P_p(|C| = n)$. As in Chapters 5 and 6, the upper bound of the last section for the radius of the finite open cluster at the origin provides a weak upper bound for $P_p(|C| = n)$. It is easy to deduce from Theorems (8.18) and (8.21) that

$$P_p\big(n \leq |C| < \infty\big) \leq \exp\big(-\eta(p)n^{1/d}\big) \qquad \text{for all } n, \tag{8.64}$$

for some $\eta(p)$ which is strictly positive when $p > p_c$, and which is uniformly bounded away from 0 on any closed sub-interval $[\alpha, \beta]$ of $(p_c, 1)$. We note that, for $d = 2$, the orders of the exponents in the lower bound (8.62) and the upper bound (8.64) are both $n^{1/2}$.

Here is a better upper bound for the tail of the size of the finite cluster at the origin.

(8.65) Theorem. *If $p > p_c$, there exists $\eta(p) > 0$ such that*

$$P_p(|C| = n) \leq \exp\big(-\eta(p)n^{(d-1)/d}\big) \qquad \textit{for all } n. \tag{8.66}$$

This provides an upper bound for the tail probabilities having the same order (in the exponent) as that of the lower bound (8.62). When $d = 2$, (8.66) is a consequence of (8.64).

Proof of Theorem (8.61). We follow Aizenman, Delyon, and Souillard (1980). Here is the idea of the proof. When $p > p_c$, each vertex has a positive probability $\theta(p)$ of belonging to the infinite open cluster. Therefore, in any specified box $B(m)$, say, the mean proportion of vertices which are in the infinite open cluster equals $\theta(p)$. In particular, the probability $\nu(m)$ that more than $\frac{1}{2}\theta(p)$ of the

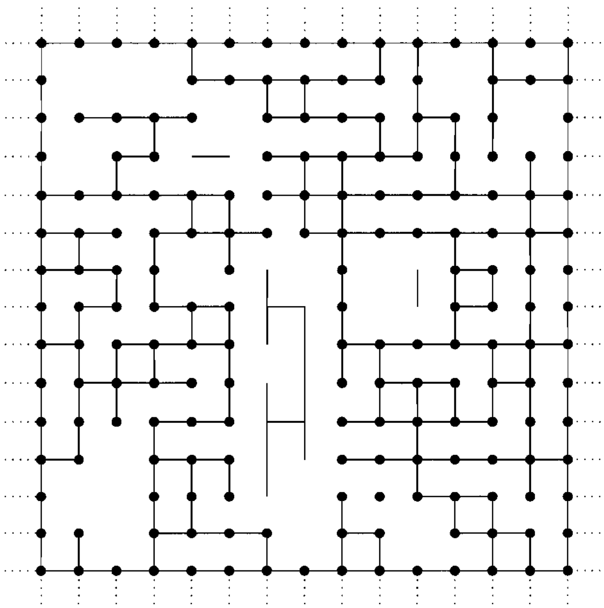


Figure 8.5. We may form a large finite open cluster by declaring all edges in $\partial B(m)$ to be open and all edges joining $\partial B(m)$ to $\partial B(m+1)$ to be closed. The cluster in question contains all vertices in $B(m)$ which are joined to $\partial B(m)$ by open paths.

vertices in $B(m)$ are joined to the surface of $B(m)$ remains bounded away from 0 as $m \to \infty$. Suppose that every edge joining two vertices in $\partial B(m)$ is open but every edge joining a vertex in $\partial B(m)$ to a vertex in $\partial B(m+1)$ is closed; the probability of this is approximately $\pi^{|\partial B(m)|}$ for some $\pi = \pi(p)$ satisfying $0 < \pi < 1$. If both of the above events occur then $B(m+1)$ contains a finite open cluster of $\mathbb{L}^d$ of size greater than $\frac{1}{2}|B(m)|\theta(p)$, and the probability of this is roughly at least $\nu(m)\pi^{|\partial B(m)|}$. See Figure 8.5 for an illustration of these events. The origin belongs to this cluster with probability approximately $\theta(p)$, so that

$$P_p\left(\tfrac{1}{2}|B(m)|\theta(p) \le |C| < \infty\right) \ge A\pi^{|\partial B(m)|} \tag{8.67}$$

for some positive constant $A = A(p, d)$. We write $n = \frac{1}{2}|B(m)|\theta(p)$, so that $|\partial B(m)|$ is of order $n^{(d-1)/d}$, and we obtain an inequality of the desired form. We shall now make this argument rigorous.

Suppose that $p_c < p < 1$ and let m be a positive integer. Let R_m be the number of vertices of $B(m)$ which lie in the (a.s. unique) infinite open cluster.

(8.68) Lemma.

$$P_p\big(R_m \geq \tfrac{1}{2}\theta(p)|B(m)|\big) \geq \tfrac{1}{2}\theta(p).$$

Proof. Clearly $R_m \leq |B(m)|$, and so

$$\begin{aligned} E_p(R_m) &\leq |B(m)|P_p\big(R_m \geq \tfrac{1}{2}\theta(p)|B(m)|\big) \\ &\qquad\qquad + \tfrac{1}{2}\theta(p)|B(m)|P_p\big(R_m < \tfrac{1}{2}\theta(p)|B(m)|\big) \\ &\leq |B(m)|P_p\big(R_m \geq \tfrac{1}{2}\theta(p)|B(m)|\big) + \tfrac{1}{2}\theta(p)|B(m)|. \end{aligned}$$

We use the fact that $E_p(R_m) = \theta(p)|B(m)|$ and divide by $|B(m)|$ to obtain the conclusion of the lemma. □

Let U_m be the number of vertices inside $B(m)$ which are joined to vertices in $\partial B(m)$ by open paths. Certainly $U_m \geq R_m$, and so

$$P_p\big(\tfrac{1}{2}\theta(p)|B(m)| \leq U_m \leq |B(m)|\big) \geq \tfrac{1}{2}\theta(p).$$

The random variable U_m is independent of the states of all edges between pairs of vertices in $\{x \in \mathbb{Z}^d : \|x\| \geq m\}$; remember that $\|x\| = \max\{|x_i| : 1 \leq i \leq d\}$. Let A_m be the event that all edges joining pairs of vertices in $\partial B(m)$ are open and all edges joining vertices of $\partial B(m)$ to vertices of $\partial B(m+1)$ are closed. It is easy to check that there exists $\pi = \pi(p)$ satisfying $0 < \pi < 1$ such that

(8.69) $$P_p(A_m) \geq \pi^{|\partial B(m)|} \qquad \text{for all } m \geq 1.$$

Furthermore, U_m is independent of A_m, so that

(8.70) $$P_p\big(A_m,\ \tfrac{1}{2}\theta(p)|B(m)| \leq U_m \leq |B(m)|\big) \geq \tfrac{1}{2}\theta(p)\pi^{|\partial B(m)|}.$$

If the event on the left side here occurs, then the vertex $(m, m, \ldots, m)$ is contained in a finite cluster of $\mathbb{L}^d$ having between $\frac{1}{2}\theta(p)|B(m)|$ and $|B(m)|$ vertices; we translate the vertex $(m, m, \ldots, m)$ to the origin to deduce that

(8.71) $$P_p\big(\tfrac{1}{2}\theta(p)|B(m)| \leq |C| \leq |B(m)|\big) \geq \tfrac{1}{2}\theta(p)\pi^{|\partial B(m)|}.$$

It is easy to deduce from this that there exists $\eta = \eta(p) < \infty$ such that

$$P_p\big(n \leq |C| < \infty\big) \geq \exp\big(-\eta n^{(d-1)/d}\big) \qquad \text{for all } n,$$

but the theorem requires more than this.

(8.72) Lemma. *Let $p > p_c$. There exists $\eta = \eta(p) < \infty$ such that the following holds. For every $i \geq 1$, there exists a positive integer $r(i)$ such that*

$$r(1) = 1, \tag{8.73}$$

$$2 \leq \frac{r(i+1)}{r(i)} \leq \delta, \tag{8.74}$$

$$P_p\big(|C| = r(i)\big) \geq r(i)\exp\big(-\eta r(i)^{(d-1)/d}\big), \tag{8.75}$$

where $\delta = 2^{d+3}\theta(p)^{-2}$.

Proof. Remember that

$$|B(m)| = (2m+1)^d, \qquad |\partial B(m)| \leq 2d(2m+1)^{d-1}. \tag{8.76}$$

From (8.71), for each $m \geq 1$, there exists $k(m)$ such that

$$\tfrac{1}{2}\theta(p)|B(m)| \leq k(m) \leq |B(m)| \tag{8.77}$$

and

$$\begin{aligned} P_p\big(|C| = k(m)\big) &\geq \frac{1}{2|B(m)|}\theta(p)\pi^{|\partial B(m)|} \\ &= \frac{1}{2|B(m)|}\theta(p)\exp\big(-|\partial B(m)|\log(1/\pi)\big). \end{aligned} \tag{8.78}$$

From (8.76) and (8.77), there exists $\nu = \nu(p) < \infty$ such that

$$|\partial B(m)|\log(1/\pi) \leq \nu k(m)^{(d-1)/d} \qquad \text{for all } m.$$

Thus, from (8.77) and (8.78),

$$P_p\big(|C| = k(m)\big) \geq \frac{\theta(p)^2}{2k(m)}\exp\big(-\nu k(m)^{(d-1)/d}\big).$$

This inequality is interesting only when $k(m)$ is large, in which case the exponential term is dominating; therefore, there exists $\eta = \eta(p) < \infty$ for which

$$P_p\big(|C| = k(m)\big) \geq k(m)\exp\big(-\eta k(m)^{(d-1)/d}\big), \tag{8.79}$$

and

$$P_p(|C| = 1) \geq e^{-\eta}. \tag{8.80}$$

From the sequence $(1, k(1), k(2), \dots)$ we remove a subsequence $(r(i) : i \geq 1)$ in the following manner. We set $r(1) = 1$. Having chosen $r(1), r(2), \dots, r(i)$ with $r(i) = k(j)$ say, we let l be the smallest integer such that

$$|B(j)| \leq \tfrac{1}{4}\theta(p)|B(l)| \tag{8.81}$$

and then we set $r(i+1) = k(l)$. We have from (8.77) that

$$k(l) \geq \tfrac{1}{2}\theta(p)|B(l)| \geq 2|B(j)| \geq 2k(j)$$

so that $2 \leq r(i+1)/r(i)$. On the other hand, l is the smallest integer such that

$$(2j+1)^d \leq \tfrac{1}{4}\theta(p)(2l+1)^d$$

so that

$$(2j+1)^d > \tfrac{1}{4}\theta(p)(2l-1)^d,$$

which is to say that

$$(2l+1)^d \leq \left\{\left(\frac{4}{\theta(p)}\right)^{1/d}(2j+1)+2\right\}^d \leq \frac{2^{d+2}}{\theta(p)}(2j+1)^d.$$

Hence, by (8.77),

$$k(l) \leq |B(l)| \leq \frac{2^{d+2}}{\theta(p)}|B(j)| \leq \frac{2^{d+3}}{\theta(p)^2}k(j),$$

giving that $r(i+1)/r(i) \leq \delta$ as required. □

We show next that any positive integer n may be constructed using the $r(i)$ as building blocks.

(8.82) Lemma. *Let δ, η, and $(r(i) : i \geq 1)$ be as given in Lemma* (8.72). *Any positive integer n may be expressed in the form*

$$n = \sum_{i=1}^{I} w_i r(i), \tag{8.83}$$

where $I = \max\{i : r(i) \leq n\}$ and the w_i are integers which satisfy

$$0 \leq w_i \leq \delta \qquad \textit{for all } i. \tag{8.84}$$

Furthermore, with such a representation,

$$\sum_{i=1}^{I} w_i r(i)^{(d-1)/d} \leq 4\delta n^{(d-1)/d}. \tag{8.85}$$

Proof. Certainly $n = 1$ can be expressed in the form (8.83) subject to (8.84); just take $w_I = I = 1$. Suppose that any integer not exceeding k can be expressed in the form (8.83) subject to (8.84), and consider the integer $n = k + 1$. Let $I = \max\{i : r(i) \le k + 1\}$. Then

$$k + 1 = r(I) + \big(k + 1 - r(I)\big) = r(I) + \sum_{i=1}^{I} w_i r(i)$$

for some $(w_i : i \ge 1)$ satisfying (8.84), from the induction hypothesis applied to $k + 1 - r(I)$. It remains to check that $1 + w_I \le \delta$. From the definition of I, we have that $r(I) \le k + 1 < r(I + 1)$. On the other hand, $r(I + 1) \le \delta r(I)$ by (8.74), giving that

$$(1 + w_I)r(I) \le k + 1 < \delta r(I)$$

as required.

Finally, suppose that n has been expressed in the form (8.83) subject to (8.84). We will show that (8.85) holds. We have from (8.84) that $w_i \le \delta$, and from (8.74) that

$$r(i) \le \tfrac{1}{2} r(i + 1) \le \frac{1}{2^{I-i}} r(I) \le \frac{n}{2^{I-i}}$$

whenever $1 \le i \le I$. Therefore,

$$\begin{aligned}
\sum_{i=1}^{I} w_i r(i)^{(d-1)/d} &\le \delta n^{(d-1)/d} \sum_{i=1}^{I} \left(\tfrac{1}{2}\right)^{(I-i)(d-1)/d} \\
&\le \delta n^{(d-1)/d} \sum_{i=0}^{\infty} \left(\tfrac{1}{2}\right)^{i/2} \qquad \text{since } d \ge 2 \\
&\le 4\delta n^{(d-1)/d}
\end{aligned}$$

as required. □

We may use the $r(i)$ as building blocks in order to complete the proof. Let n be a positive integer, and write n in the form

(8.86) $$n = \sum_{i=1}^{I} w_i r(i) \qquad \text{where } 0 \le w_i \le \delta,$$

as in Lemma (8.82). We recall Lemma (6.102), which asserts that

$$\frac{1}{r+s} P_p(|C| = r + s) \ge pq^{-2} \frac{1}{r} P_p(|C| = r) \frac{1}{s} P_p(|C| = s) \qquad \text{if } r, s \ge 1.$$

The proof of this is valid whenever $0 < p = 1 - q < 1$. It follows from (8.86) that

(8.87) $$\frac{1}{n} P_p(|C| = n) \ge (pq^{-2})^w \prod_{i=1}^{I} \left\{ \frac{1}{r(i)} P_p\big(|C| = r(i)\big) \right\}^{w_i},$$

where $w = \sum_{i=1}^{I} w_i$. However, $w_i \le \delta$ for all i, and $r(i) \ge 2^{i-1}$ by (8.74), giving that

$$I = \max\{i : r(i) \le n\} \le 1 + \log_2 n$$

and

$$w \le \delta(1 + \log_2 n).$$

Writing $\zeta = \min\{1, pq^{-2}\}$, we have from (8.75) and (8.85) that

$$\begin{aligned}\frac{1}{n} P_p(|C| = n) &\ge \zeta^{\delta(1+\log_2 n)} \exp\Big(-\eta \sum_{i=1}^{I} w_i r(i)^{(d-1)/d}\Big)\\ &\ge \zeta^{\delta(1+\log_2 n)} \exp\big(-4\delta\eta n^{(d-1)/d}\big)\end{aligned}$$

and we are home. □

Proof of Theorem (8.65). We follow the argument of Kesten and Zhang (1990), and may suppose that $d \ge 3$. The first step is to show such an inequality for values of p which are sufficiently close to 1 (this was proved by Kunz and Souillard (1978); see also Kesten (1982, p. 99)). The idea is the following. There exists $\nu > 0$ such that, if the origin is in a finite cluster C of size n, there are at least $\nu n^{(d-1)/d}$ closed edges in the external edge boundary of C; see Figure 1.6 for a drawing of the external edge boundary of C in two dimensions. Furthermore, at least one such edge has an endvertex within distance n of the origin. Let $\mathbb{L}_c^d$ denote the covering lattice of $\mathbb{L}^d$ (see Section 1.6). We define a new 'neighbour relation' on the set of edges of $\mathbb{L}^d$, with the effect that the edges in the external edge boundary of C correspond to the vertex set of a connected subgraph of $\mathbb{L}_c^d$. We now use the fact that the number of such connected subgraphs having m vertices and containing a specified vertex is at most $\mu(d)^m$, for some constant $\mu(d)$ depending only on the number of dimensions. The usual 'Peierls argument' gives

(8.88) $$P_p(|C| = n) \le N_n \big\{q\mu(d)\big\}^{\nu n^{(d-1)/d}},$$

where $p+q = 1$ and $N_n = O(n^d)$ is the number of edges of $\mathbb{L}^d$ having an endvertex within distance n of the origin. If p is sufficiently close to 1, then $q\mu(d) < 1$ and the result is proved. It is rather easy to make this argument rigorous in two dimensions, but there are certain topological complications when $d \ge 3$, and we omit these details. Before moving on, we note that the above argument may in principle be adapted to bond and site percolation on any lattice, and not merely bond percolation on $\mathbb{L}^d$. This amounts to a proof of the following fact, which we shall use in the next stage of the proof. There exists $p' \in (p_c^{\text{site}}(\mathbb{L}^d), 1)$ and $\eta(p)$ satisfying $\eta(p) > 0$ when $p > p'$, such that (8.66) holds for site percolation on $\mathbb{L}^d$.

The proof of the theorem continues as follows. Let N be a positive integer. From the lattice $\mathbb{L}^d$ we construct a 'renormalized' lattice $\mathcal{R}_N(\mathbb{L}^d)$ as follows.

$\mathcal{R}_N(\mathbb{L}^d)$ has $\mathbb{Z}^d$ as vertex set, the vector $x \in \mathbb{Z}^d$ corresponding to the box

(8.89) $$B_x = B_x(N) = \prod_{i=1}^{d} \big[x_i N, (x_i + 1)N\big)$$

of $\mathbb{L}^d$; we declare two vertices x and y of $\mathcal{R}_N(\mathbb{L}^d)$ to be *adjacent* if B_x and B_y have a common 'hyperface' in $\mathbb{L}^d$, and as a consequence $\mathcal{R}_N(\mathbb{L}^d)$ is isomorphic to $\mathbb{L}^d$. We declare a vertex x of $\mathcal{R}_N(\mathbb{L}^d)$ to be *bad* if either:

(a) there exists no open path of $\mathbb{L}^d$ from B_x to the surface $\partial B'_x$ of $B'_x = \prod_{i=1}^{d}\big[(x_i - 2)N, (x_i + 3)N\big)$; or

(b) there exists a neighbour y of x in $\mathcal{R}_N(\mathbb{L}^d)$ such that there exist open paths of $\mathbb{L}^d$ from B_x to $\partial B'_x$ and from B_y to $\partial B'_x$ with the property that these paths lie in disjoint open clusters of B'_x.

If neither (a) nor (b) holds, we call x *good.*

The family $X = \{X_x : x \in \mathbb{Z}^d\}$ given by

$$X_x = X_x(N) = \begin{cases} 1 & \text{if } x \text{ is good,} \\ 0 & \text{otherwise.} \end{cases}$$

is k-dependent for some integer k which does not vary with the choice of N (recall the discussion of k-dependence in Section 7.4).

Since $\theta(p) > 0$,

$$P_p(B_x \leftrightarrow \infty) \to 1 \quad \text{as } N \to \infty,$$

and it follows by Lemma (7.89) that

(8.90) $$P_p(x \text{ is good}) \to 1 \quad \text{as } N \to \infty, \ \text{for } x \in \mathbb{Z}^d.$$

By Theorem (7.65), we may find a real number π satisfying $\pi > p'$, and an integer N, such that X dominates (stochastically) a site percolation process on $\mathbb{L}^d$ with (supercritical) density π.

Let C be the open cluster of $\mathbb{L}^d$ at the origin, as usual. We define the *external* $\mathcal{R}_N(\mathbb{L}^d)$*-boundary* of C to be the set of vertices x of $\mathcal{R}_N(\mathbb{L}^d)$ such that:

(a) $B_x \cap C = \varnothing$;

(b) x has a neighbour y such that $B_y \cap C \neq \varnothing$; and

(c) there exists an infinite path (not necessarily open) of $\mathbb{L}^d$ which has an end-vertex in B_x and which is disjoint from C.

Suppose that $|C| = n$. By the isoperimetric inequality, the external $\mathcal{R}_N(\mathbb{L}^d)$-boundary of C has order at least $n^{(d-1)/d}$. Furthermore, if n is sufficiently large compared with N, it is not difficult to see that every vertex in the external $\mathcal{R}_N(\mathbb{L}^d)$-boundary of C is bad. We may adapt the proof leading to (8.88) to obtain that the probability of such a closed 'surface' surrounding the origin of $\mathcal{R}_N(\mathbb{L}^d)$ is at most some exponentially decaying function of its order $n^{(d-1)/d}$. □

8.7 Differentiability of θ, χ^{f}, and κ

Once we have a decent upper bound for $P_p(n \le |C| < \infty)$ it is easy to show the infinite differentiability of macroscopic quantities such as the percolation probability $\theta(p)$, the mean size $\chi^{\mathrm{f}}(p)$ of finite open clusters, and the number $\kappa(p)$ of open clusters per vertex. It is not necessary to have a really fancy bound for $P_p(n \le |C| < \infty)$; anything which decreases faster than $n^{-\alpha}$ for all $\alpha > 0$ will suffice. We will use here the bound (8.64):

(8.91) $$P_p\big(n \le |C| < \infty\big) \le \exp\big(-\eta(p)n^{1/d}\big) \qquad \text{for all } n,$$

where $\eta(p)$ is uniformly bounded away from 0 on any closed sub-interval of $(p_{\mathrm{c}}, 1)$.

(8.92) Theorem. Differentiability of θ, χ^{f}, and κ above p_{c}. *The percolation probability θ, the mean size χ^{f} of the finite open cluster at the origin, and the number κ of open clusters per vertex are infinitely differentiable functions of p on the interval $(p_{\mathrm{c}}, 1]$.*

The proof shows that $E_p(f(|C|); |C| < \infty)$ is an infinitely differentiable function of p on $(p_{\mathrm{c}}, 1]$, for any polynomial f in $|C|$ and $|C|^{-1}$. It is an open question of debatable interest to decide whether or not such functions are in general *analytic* on $(p_{\mathrm{c}}, 1]$.

Proof. We follow the argument of Russo (1978). We shall show that θ is infinitely differentiable on $(p_{\mathrm{c}}, 1]$, and the same argument is valid for χ^{f} and κ. Suppose that $p_{\mathrm{c}} < p = 1 - q \le 1$.

We write θ in the form

(8.93) $$\theta(p) = 1 - \sum_{n=1}^{\infty} P_p(|C| = n),$$

where

(8.94) $$P_p(|C| = n) = \sum_{m,b} a_{nmb}\, p^m q^b,$$

and a_{nmb} is the number of animals with n vertices, m edges, and b boundary edges. In order to show that θ is differentiable k times, we need to show the uniform convergence of

(8.95) $$\sum_{n=N}^{\infty} \frac{d^k}{dp^k} P_p(|C| = n) = \sum_{n=N}^{\infty} \sum_{m,b} a_{nmb} \frac{d^k}{dp^k}\,(p^m q^b)$$

as $N \to \infty$, on suitable sub-intervals of $(p_{\mathrm{c}}, 1]$.

Suppose first that $p_{\mathrm{c}} < p < 1$ and choose α and β such that $p_{\mathrm{c}} < \alpha < p < \beta < 1$. We shall show that the series in (8.95) is uniformly convergent on $[\alpha, \beta]$. For any $k \geq 0$, it is the case that

$$\begin{aligned}\left|\frac{d^k}{dp^k}(p^m q^b)\right| &= \left|\sum_{r=0}^{k}\binom{k}{r} m_r b_{k-r} p^{m-r}(-1)^{k-r} q^{b-(k-r)}\right| \\ &\leq p^m q^b \sum_{r=0}^{k}\binom{k}{r}(m/p)^r (b/q)^{k-r} \\ &= p^m q^b \left(\frac{m}{p}+\frac{b}{q}\right)^k\end{aligned}$$

where $x_r = x(x-1)\dots(x-r+1)$. We shall use the fact that $a_{nmb} = 0$ unless $b \leq 2dn$ and $m \leq dn$, and may therefore assume that

$$\frac{m}{p}+\frac{b}{q} \leq \frac{2dn}{pq}.$$

By (8.95),

$$\left|\sum_{n=N}^{\infty}\frac{d^k}{dp^k}P_p(|C|=n)\right| \leq \left(\frac{2d}{pq}\right)^k \sum_{n=N}^{\infty} n^k P_p(|C|=n). \tag{8.96}$$

By (8.91),

$$P_p(|C|=n) \leq \exp\big(-\widehat{\eta} n^{1/d}\big) \qquad \text{for all } n,$$

uniformly in $p \in [\alpha, \beta]$, where $\widehat{\eta} = \inf\{\eta(p) : \alpha \leq p \leq \beta\} > 0$. Therefore, the sum in (8.96) is uniformly convergent on $[\alpha, \beta]$, giving that θ is k times differentiable at p.

Showing (one-sided) differentiability of θ at $p = 1$ seems to require a certain amount of topology which the author is eager to avoid. We use (8.88) instead of (8.91) in order to bound a typical term in (8.96). This gives the following upper bound for the series in (8.96) when $p < 1$:

$$\left(\frac{2d\mu(d)}{p}\right)^k \sum_{n=N}^{\infty} n^k N_n \{q\mu(d)\}^{-k+\nu n^{(d-1)/d}},$$

where $N_n \leq A_1 n^d$ for some constant A_1, $1 < \mu(d) < \infty$, and $\nu > 0$. If $0 < q = 1 - p \leq \frac{1}{2}\mu(d)^{-1}$ say, then (8.96) is bounded above by

$$A_2 \sum_{n=N}^{\infty} n^{k+d} 2^{-\nu' n^{(d-1)/d}}$$

for some constant A_2 and some $\nu' > 0$. This bound does not depend on p, and it is easy to deduce that θ is k times differentiable (from the left) at $p = 1$. □

8.8 Geometry of the Infinite Open Cluster

When $\theta(p) > 0$, there exists almost surely a unique infinite open cluster. What does this cluster look like? We are particularly interested in answering this question when p is only slightly greater than p_c, since the infinite open cluster has only low density in this case. Indeed, if $\theta(p_c) = 0$ as expected, then the density of the infinite open cluster decreases to 0 as p tends to p_c from above.

We study only two aspects of the geometry of the infinite open cluster. First, we shall show that, with probability approaching 1 as $n \to \infty$, there exists an open path traversing the box $B(n)$ when n is large. Secondly, we shall see that the 'surface' of the infinite cluster has the same order as its volume. An interesting facet of the two theorems following is that they use the hypothesis that $\theta(p) > 0$ rather than that $p > p_c$. No quantitative argument is employed.

Let $B(n)$ be the box with side-length $2n$ and centre at the origin. A *left–right crossing* of $B(n)$ is an open path of $B(n)$ joining some vertex x with $x_1 = -n$ to some vertex y with $y_1 = n$; we denote by $\mathrm{LR}(n)$ the event that there exists a left–right crossing of $B(n)$. It is clear that $P_p(\mathrm{LR}(n))$ decays exponentially to 0 as $n \to \infty$ when $0 < p < p_c$, since $\mathrm{LR}(n)$ occurs only if one of the $(2n+1)^{d-1}$ vertices on the left face of $B(n)$ belongs to an open cluster of size at least $2n + 1$. The picture is quite different when $p > p_c$, since there exists an infinite open cluster in this case.

(8.97) Theorem. *If $\theta(p) > 0$ then $P_p(\mathrm{LR}(n)) \to 1$ as $n \to \infty$.*

The proof of this result uses only the uniqueness of the infinite open cluster when $\theta(p) > 0$, and it is for this reason that the theorem contains no estimate for the rate of convergence to 1 of $P_p(\mathrm{LR}(n))$. Actually, this convergence takes place extremely quickly when $p > p_c$ in that

$$P_p\big(\mathrm{LR}(n)\big) \geq 1 - \exp\big(-\beta(p)n^{d-1}\big) \tag{8.98}$$

for all n and some β satisfying $\beta(p) > 0$; this form of the exponent is motivated by the fact that n^{d-1} is the order of the number of possible endvertices of left–right crossings of $B(n)$. This and more was proved in Theorem (7.68), using much deeper methods.

We turn now to the ratio of boundary to volume of the infinite cluster. Let I be the set of vertices of $\mathbb{L}^d$ which belong to infinite open clusters; we write I_e for the set of open edges which join pairs of vertices in I, and ΔI for the set of closed edges of $\mathbb{L}^d$ which have at least one endvertex in I. We express the boundary/volume ratio of I as

$$\lim_{n\to\infty} \frac{|\Delta I \cap B(n)|}{|I_e \cap B(n)|},$$

where $I_e \cap B(n)$ and $\Delta I \cap B(n)$ are interpreted as the subsets of I_e and ΔI containing edges which join pairs of vertices in $B(n)$.

(8.99) Theorem. *If $\theta(p) > 0$ then, as $n \to \infty$,*

$$\frac{|\Delta I \cap B(n)|}{|I_{\mathrm{e}} \cap B(n)|} \to \frac{1-p}{p} \qquad \textit{almost surely.}$$

Proof of Theorem (8.97). Suppose that $\theta(p) > 0$, so that there exists almost surely a unique infinite open cluster. Let ε be small and positive, and choose m large enough so that

$$P_p\big(I(m)\big) > 1 - \varepsilon, \tag{8.100}$$

where $I(m)$ is the event that some vertex of $B(m)$ lies in the infinite open cluster. Let n be an integer satisfying $n > m$, and let $F_1, F_2, \dots, F_{2d}$ be a list of the $2d$ faces of $B(n)$. If $I(m)$ occurs then some vertex of $B(m)$ is joined by an open path to some face F_i. Thus

$$\begin{aligned} 1 - P_p\big(I(m)\big) &\geq 1 - P_p\left(\bigcup_{i=1}^{2d}\big\{B(m) \leftrightarrow F_i \text{ in } B(n)\big\}\right) \\ &= P_p\left(\bigcap_{i=1}^{2d}\big\{B(m) \leftrightarrow F_i \text{ in } B(n)\big\}^{\mathrm{c}}\right). \end{aligned}$$

The last $2d$ events are decreasing events having equal probability. We apply the FKG inequality to deduce that

$$1 - P_p\big(I(m)\big) \geq \Big\{1 - P_p\big(B(m) \leftrightarrow F \text{ in } B(n)\big)\Big\}^{2d}$$

for any given face F of $B(n)$. When combined with (8.100), this implies that

$$P_p\big(B(m) \leftrightarrow F \text{ in } B(n)\big) \geq 1 - \varepsilon^{1/(2d)}. \tag{8.101}$$

Writing $F_{\mathrm{L}}(n)$ and $F_{\mathrm{R}}(n)$ for the left and right faces of $B(n)$ (that is, the sets $\{x \in \partial B(n) : x_1 = -n\}$ and $\{x \in \partial B(n) : x_1 = n\}$) we have by the FKG inequality that

$$\begin{aligned} P_p\Big(\big\{B(m) \leftrightarrow F_{\mathrm{L}}(n) \text{ in } B(n)\big\} \cap \big\{B(m) \leftrightarrow F_{\mathrm{R}}(n) \text{ in } B(n)\big\}\Big) \\ \geq \big\{1 - \varepsilon^{1/(2d)}\big\}^2. \end{aligned} \tag{8.102}$$

See Figure 8.6 for an illustration of this event. Two open paths from $B(m)$ to $F_{\mathrm{L}}(n)$ and to $F_{\mathrm{R}}(n)$ may be combined to form a left–right crossing of $B(n)$ whenever they are connected in $B(n)$ by an open path (see the figure again). The chance that there exist two such paths which are not connected by an open path of $B(n)$ is no greater than the probability of the event

$$A_{m,n} = \big\{\text{there exist two vertices of } \partial B(m) \text{ which are in disjoint open clusters of } B(n) \text{ both of which intersect } \partial B(n)\big\};$$

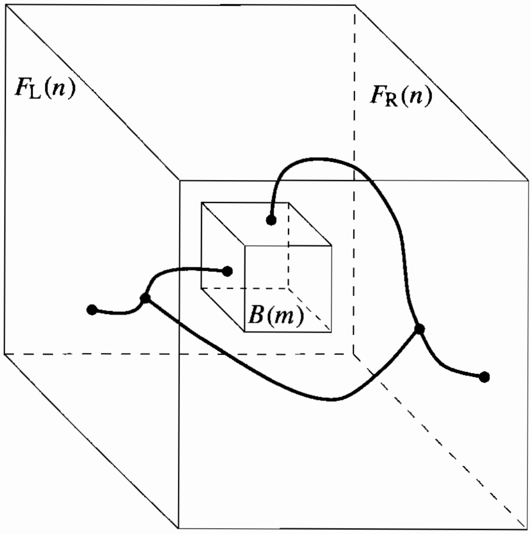


Figure 8.6. The smaller box is $B(m)$ and the larger is $B(n)$. The sets $F_L(n)$ and $F_R(n)$ are the 'left' and 'right' faces of $B(n)$. This is a sketch of open paths joining $B(m)$ to $F_L(n)$ and to $F_R(n)$, together with an open path of $B(n)$ linking the first two paths.

however, $A_{m,n} \supseteq A_{m,n+1}$, and it follows that

$$\begin{aligned} P_p(A_{m,n}) \to P_p\big(&\text{there exist two vertices of } \partial B(m) \text{ in disjoint infinite} \\ &\text{open clusters}\big) \\ = 0& \end{aligned}$$

as $n \to \infty$, by the uniqueness theorem for the infinite open cluster. Thus, by (8.102),

$$\begin{aligned} P_p\big(\mathrm{LR}(n)\big) &\geq \big\{1 - \varepsilon^{1/(2d)}\big\}^2 - P_p(A_{m,n}) \\ &\to \big\{1 - \varepsilon^{1/(2d)}\big\}^2 \qquad \text{as } n \to \infty. \end{aligned}$$

We take the limit as $\varepsilon \downarrow 0$ to obtain the required conclusion. □

Proof of Theorem (8.99). We follow Grimmett (1981b). Suppose $\theta(p) > 0$. For each edge e of $\mathbb{L}^d$, we let χ_e and β_e be the indicator functions of the events that e belongs to I_e and ΔI, respectively. The families $(\chi_e : e \in \mathbb{E}^d)$ and $(\beta_e : e \in \mathbb{E}^d)$ are stationary under translations of the lattice; furthermore, all events which are invariant under such translations have probability either 0 or 1. By the ergodic theorem (see Dunford and Schwartz (1958, Theorem VIII.6.9) or Tempel'man (1972, Theorem 6.1 and Corollary 6.2)) we have that, as $n \to \infty$,

$$\frac{1}{|\mathbb{E}_n|} \sum_{e \in \mathbb{E}_n} \chi_e \to E_p(\chi_e), \qquad \frac{1}{|\mathbb{E}_n|} \sum_{e \in \mathbb{E}_n} \beta_e \to E_p(\beta_e),$$

almost surely, where $\mathbb{E}_n$ is the set of edges having both endvertices in $B(n)$. However,

$$E_p(\chi_e) = P_p\big(e \text{ belongs to an infinite open cluster}\big)$$

and

$$\begin{aligned} E_p(\beta_e) &= P_p\big(e \text{ closed, some endvertex of } e \text{ is in an infinite open cluster}\big) \\ &= q P_p^e\big(\text{some endvertex of } e \text{ is in an infinite open cluster}\big) \\ &= qp^{-1} P_p\big(e \text{ open, some endvertex of } e \text{ is in an infinite open cluster}\big) \\ &= qp^{-1} P_p\big(e \text{ is in an infinite open cluster}\big), \end{aligned}$$

where P_p^e is the probability measure of percolation on $\mathbb{L}^d$ with e removed, and $p + q = 1$. Thus, as $n \to \infty$,

$$\begin{aligned} \frac{|\Delta I \cap B(n)|}{|I_e \cap B(n)|} &= \sum_{e \in \mathbb{E}_n} \beta_e \bigg/ \sum_{e \in \mathbb{E}_n} \chi_e \\ &\to qp^{-1} \qquad \text{almost surely.} \end{aligned}$$

□

8.9 Notes

Section 8.2. The uniqueness of the infinite open cluster was proved by Aizenman, Kesten, and Newman (1987a, b); Newman and Schulman (1981a, b) deal with related material. The proof was simplified by Gandolfi, Grimmett, and Russo (1988), and further generalized by Gandolfi (1989) to yield a proof that there exists at most one infinite cluster for any stationary finite-range Gibbs state in any dimension. A similar conclusion was reached by Gandolfi, Keane, and Russo (1988) for any ergodic translation-invariant measure on the state space $\{0, 1\}^{\mathbb{E}^2}$ in two dimensions, whenever the FKG inequality remains in force. The proof of uniqueness given here is a beautiful advance by Burton and Keane (1989). Unlike earlier proofs, it uses no explicit estimate, and is therefore well suited for adaptation to other settings, where it has indeed been useful (see, for example, Bezuidenhout and Grimmett (1997)).

It has been asked whether or not the infinite open cluster is unique 'simultaneously' for all values of p satisfying $\theta(p) > 0$. Let η_p be given by (1.4), and let N_p be the number of infinite open clusters in the configuration η_p. Alexander (1995a) has proved that

$$P\big(N_p \in \{0, 1\} \text{ for all } p\big) = 1.$$

As for the number of infinite open clusters of bond percolation on *subgraphs* of $\mathbb{L}^d$, Barsky, Grimmett, and Newman (1991a, b) have proved the uniqueness theorem for certain graphs including $\mathbb{Z}_+^e \times \mathbb{Z}^{d-e}$ for $1 \le e \le d$. Other results in this direction are due to Aizenman, Chayes, Chayes, Fröhlich, and Russo (1983), Kesten (1988b, 1992), and Newman and Wu (1988).

The situation concerning the number N of infinite open clusters is, of course, quite different for percolation on a tree rather than on a lattice (see Section 10.1), and it is elementary that (almost surely) $N = \infty$ for such a graph whenever $N \ge 1$. For general periodic graphs G, the proof of uniqueness may be adapted successfully so long as the surface/volume ratio of the ball with radius n tends to 0 as $n \to \infty$. For graphs G which do not satisfy this condition, there exist generally three 'phases', corresponding to $N = 0$, $N = 1$, and $N = \infty$. See Grimmett and Newman (1990) and Zhang (1996b) for examples of such results. Such observations have provoked related work for the contact model on the tree; see, for example, Pemantle (1992), Liggett (1996a, b), and Stacey (1996).

Häggström and Peres (1998) have shown for a wide class of graphs that the property of 'having a.s. a unique infinite open cluster' is increasing in p; see also Schonmann (1998a, b). These authors study percolation on rather general periodic graphs, not merely finite-dimensional lattices.

Section 8.3. The right-continuity of θ was observed by Russo (1978), whilst the relationship between its left-continuity and the uniqueness of the infinite open cluster was discovered by van den Berg and Keane (1984). Results of Harris (1960) and Kesten (1980a) imply that $\theta(p_c(2)) = 0$ in two dimensions. Aizenman, Kesten, and Newman (1987a, b) noted the continuity of the connectivity functions. Barsky, Grimmett, and Newman (1991a, b) have proved the continuity of the percolation probability for bond percolation on subgraphs such as half-spaces. Hara and Slade (1990, 1994) have shown that θ is continuous at $p_c(d)$ if d is sufficiently large, currently $d \ge 19$; see Section 10.3.

The continuity of the function

$$\upsilon(p) = P_p\big(G \text{ has a unique infinite open cluster}\big)$$

has been explored for general graphs G by Häggström and Peres (1998) and by Schonmann (1998a, b).

Section 8.4. Theorem (8.18) comes from Chayes, Chayes, Grimmett, Kesten, and Schonmann (1989). Our proof of Theorem (8.21) is a reworking of the corresponding results of Chayes, Chayes, and Newman (1987) with the word 'invasion' removed. Hammersley (1963) grew clusters recursively; see also Klein and Shamir (1982) if you can find a copy.

Section 8.5. See Chayes, Chayes, and Newman (1987) and Chayes, Chayes, Grimmett, Kesten, and Schonmann (1989) for discussions of the truncated connectivity functions and the supercritical correlation length.

Section 8.6. Aizenman, Delyon, and Souillard (1980) proved Theorem (8.61) in a more general context than percolation. Similar estimates in two dimensions have been performed by Kesten (1982, p. 98). Theorem (8.65) was proved by Kesten and Zhang (1990) subject to the hypothesis that $p_c^{slab} = p_c$; see Theorem (7.2). Steps were taken earlier in this direction by Kunz and Souillard (1978) (with loose topological ends, later secured by Kesten (1986a, Lemma 2.24)) and Chayes, Chayes, and Newman (1987).

In a recent major work, Cerf (1998b) has studied the shapes of large finite clusters in three dimensions, and has proved a general large deviation theorem. He has verified the so called Wulff shape for large finite 'droplets', and has proved the existence of the limit in (8.63) when $d = 3$. This work is related to the earlier work for the two-dimensional case by Alexander, Chayes, and Chayes (1990) and Cerf (1998a).

Section 8.7. This technique of proving differentiability appears first in Russo (1978), and later in Brånvall (1980), Grimmett (1981b), Kesten (1982), and Chayes, Chayes, and Newman (1987), and probably elsewhere also.

Section 8.8. Russo (1987, unpublished) has pointed out that the uniqueness of the infinite open cluster implies the existence of open crossings of large cubes; this argument was probably known to others also. A weaker form of Theorem (8.99) was anticipated in the physics literature (see Hankey (1978), for example), and appeared with complete proof in Newman and Schulman (1981a, b), as well as in its current form in Grimmett (1981b); the last paper contains also a bivariate central limit theorem for the pair $(|I_e \cap B(n)|, |\Delta I \cap B(n)|)$ in two dimensions.

The geometry of the infinite open cluster may be explored in depth, when $d = 2$ using path-intersection arguments, and when $d \geq 3$ using slab constructions and Theorem (7.2). Grimmett, Kesten, and Zhang (1993) have shown that random walk on the infinite open cluster is recurrent when $d = 2$, and is a.s. transient when $d \geq 3$. This work has been continued by Benjamini, Pemantle, and Peres (1998), Häggström and Mossel (1998), Levin and Peres (1998), and Hoffman and Mossel (1998). A central limit theorem for such a random walk was proved by DeMasi, Ferrari, Goldstein, and Wick (1985, 1989); see also Bezuidenhout and Grimmett (1997).

Certain large-deviation estimates have been obtained by Pisztora (1996) and Deuschel and Pisztora (1996) concerning the size of the largest open cluster inside a large box, and other quantitities.

Chapter 9

Near the Critical Point: Scaling Theory

9.1 Power Laws and Critical Exponents

The behaviour of the percolation process on $\mathbb{L}^d$ depends dramatically on whether $p < p_c$ or $p > p_c$. In the former subcritical case, all open clusters are almost surely finite and their size-distribution has a tail which decays exponentially. In the latter supercritical case, there exists almost surely an infinite open cluster and the size distribution of the remaining finite open clusters has a tail which decays slower than exponentially. Some of the major differences between these two phases are highlighted in the following table.

	Subcritical phase $p < p_c$	Supercritical phase $p > p_c$
Number of infinite open clusters	0	1
Percolation probability	$\theta(p) = 0$	$\theta(p) > 0$
Mean cluster size	$\chi(p) < \infty$	$\chi(p) = \infty$
Tail of finite clusters $P_p(n \le \|C\| < \infty)$	$\approx e^{-\zeta(p)n}$	$\ge e^{-\gamma(p)n^{(d-1)/d}}$

Table 9.1. Some differences between subcritical and supercritical percolation.

It is clear that interesting phenomena occur when p is near to its critical value p_c. Up to now, mathematicians have made great efforts but only limited progress towards understanding such phenomena, although they have been spurred on by the rich predictions and discoveries of more applied scientists. The physical theory of phase transitions and critical phenomena is well developed, and such techniques as scaling theory and renormalization are central to this theory. The mathematical foundations of such techniques and their applications remain uncertain, and it is a mathematical challenge of major importance to rectify this situation.

Renormalization and scaling theory make certain predictions about the behaviour of the percolation process when p is near or equal to its critical value p_c. With the present state of knowledge, many believe that these predictions cannot be appreciated fully without reference to the intuitive notions which motivate them. That is to say, whilst it may be possible to present simple and concrete statements about the conjectured behaviour of θ, χ, and other macroscopic quantities when p is near to p_c, such conjectures remain only superficial aspects of the scaling theory which gave birth to them. The challenge to mathematicians is therefore to *make sense of scaling theory*, rather than to verify a bunch of conjectures.

We shall nevertheless begin this chapter with a bunch of conjectures, and we hope that these conjectures will be sufficiently remarkable to capture the reader's attention for the remainder of this chapter at least. The material relating to the critical phenomenon is organized as follows. This chapter is devoted to a brief general introduction to scaling theory, including an innocent's guide to the bones of renormalization. This incorporates the idea of correlation length and the famous scaling relations. Surprisingly little of this material has yet been made rigorous; in the next chapter, we discuss the achievements and limitations of the current rigorous theory. We suppose throughout that $d \geq 2$.

The state of knowledge of critical and near-critical percolation depends greatly on whether d is large or small. Remarkable progress has been made for large d, and a rigorous and fairly full picture has emerged in recent years. In rough terms, the phase transition for large d is qualitatively similar to that of a binary tree; the interactions which arise within the lattice structure may be understood using a technique known as the 'lace expansion'. This is a long and technical story, the highlights of which are summarized in Section 10.3. The lace expansion is effective for large d only. The current requirement is that $d \geq 19$, but there is hope that the technique may be effective for $d > 6$.

We move on towards our bunch of conjectures. These concern largely the behaviour of θ, χ, χ^f, and κ when p is near to p_c. Consider first the percolation probability $\theta(p)$ when $p - p_c$ is small. It is believed that $\theta(p)$ decays to 0 in the manner of a power of $p - p_c$ as $p \downarrow p_c$. That is to say, we conjecture that, as $p \downarrow p_c$,

$$\theta(p) \approx (p - p_c)^\beta \qquad \text{for some } \beta > 0; \tag{9.1}$$

the exponent β depends of course on the value of d. It is not clear how strong we may expect such an asymptotic relation to be, and it is for this reason that we use the logarithmic relation $\approx$. Thus we conjecture that

$$\lim_{p \downarrow p_c} \frac{\log \theta(p)}{\log(p - p_c)} = \beta. \tag{9.2}$$

We note that this conjecture requires that $\theta(p_c) = 0$, which is itself an open question when $3 \le d \le 18$ (see Lemma (11.12) when $d = 2$, and Section 10.3 when $d \ge 19$).

We turn next to the mean cluster size $\chi(p)$ and the mean size $\chi^f(p)$ of a finite open cluster. It is believed that there exist $\gamma, \gamma' > 0$ such that

(9.3) $$\chi(p) \approx (p_c - p)^{-\gamma} \qquad \text{as } p \uparrow p_c$$

and

(9.4) $$\chi^f(p) \approx (p - p_c)^{-\gamma'} \qquad \text{as } p \downarrow p_c,$$

where the asymptotic relations are to be interpreted in the logarithmic manner of (9.2). It is believed further that $\gamma = \gamma'$. Of course, $\chi(p) = \chi^f(p)$ for $p < p_c$, so that these conjectures amount jointly to asserting that

(9.5) $$\chi^f(p) \approx |p - p_c|^{-\gamma} \qquad \text{as } p \to p_c$$

for some $\gamma > 0$.

It is believed that the number $\kappa(p)$ of open clusters per vertex is twice but not thrice differentiable at $p = p_c$. We thus conjecture that the third derivative of κ satisfies

(9.6) $$\kappa'''(p) \approx |p - p_c|^{-1-\alpha} \qquad \text{as } p \to p_c,$$

where $-1 \le \alpha < 0$. This is not a very natural way to express the conjecture, but there are certain historical reasons for expressing the exponent in this form.

The functions θ, χ^f, and κ are all expressible as moments of $|C|$ on the event that $|C| < \infty$:

$$\begin{aligned}
\theta(p) &= 1 - E_p(1; |C| < \infty), \\
\chi^f(p) &= E_p(|C|; |C| < \infty), \\
\kappa(p) &= E_p(|C|^{-1}; |C| < \infty),
\end{aligned}$$

and thus it is natural to introduce a critical exponent for each conditional moment of $|C|$. It is believed that there exists $\Delta > 0$ such that

(9.7) $$\frac{E_p(|C|^{k+1}; |C| < \infty)}{E_p(|C|^k; |C| < \infty)} \approx |p - p_c|^{-\Delta} \qquad \text{as } p \to p_c,$$

whenever $k \ge 1$. The hypothetical quantity Δ is called the *gap exponent*.

We move on to consider the case $p = p_c$. If $\theta(p_c) = 0$ then $\chi(p_c) = \sum_n n P_{p_c}(|C| = n)$. We saw in (6.52) that $\chi(p_c) = \infty$, and this suggests that

$P_{p_c}(|C| = n)$ decays in the manner of a negative power of n. We therefore conjecture that there exists $\delta \geq 1$ such that

$$(9.8) \qquad P_{p_c}(|C| = n) \approx n^{-1-1/\delta} \qquad \text{as } n \to \infty;$$

there are historical reasons for writing the exponent in this form. The asymptotic relation (9.8) implies that

$$(9.9) \qquad P_{p_c}(n \leq |C| < \infty) \approx n^{-1/\delta} \qquad \text{as } n \to \infty.$$

We expect similarly that the probability $P_{p_c}(0 \leftrightarrow \partial B(n))$, that there exists an open path from the origin to the surface of the box $B(n)$, decays in the manner of

$$(9.10) \qquad P_{p_c}\big(0 \leftrightarrow \partial B(n)\big) \approx n^{-1/\rho} \qquad \text{as } n \to \infty$$

for some $\rho > 0$; some authors use the notation δ_r in place of ρ. We make the slightly stronger conjecture that

$$(9.11) \qquad P_{p_c}\big(\mathrm{rad}(C) = n\big) \approx n^{-1-1/\rho} \qquad \text{as } n \to \infty;$$

as usual, the *radius* of C is defined as in (6.3) by $\mathrm{rad}(C) = \max\{|x| : x \in C\}$ where $|x| = \delta(0, x)$. We expect that the probability $\tau_{p_c}(0, x)$, that there exists an open path joining the origin to x, should decay in the manner of a negative power of $|x|$, and we conjecture therefore that

$$(9.12) \qquad \tau_{p_c}(0, x) \approx |x|^{2-d-\eta} \qquad \text{as } |x| \to \infty,$$

for some η. We summarize these and three further conjectures in Table 9.2.

These relations involve seven 'critical exponents' α, β, γ, δ, Δ, η, ρ, each of which depends on the value of d. No proofs are known of these relations for general d.

In the next section, we sketch the physical theory of 'scaling', which predicts that the seven critical exponents are not independent of each other but satisfy two sets of relations, called 'scaling' and 'hyperscaling' relations. The *scaling relations* assert that α, β, γ, δ, Δ satisfy

$$(9.13) \qquad 2 - \alpha = \gamma + 2\beta = \beta(\delta + 1),$$

$$(9.14) \qquad \Delta = \delta\beta,$$

and the validity of these equations is widely accepted. The hyperscaling relations are more questionable, and we require some preliminary discussion before we may state these relations.

A principal concept of scaling theory is the idea of 'correlation length', and we motivate this concept by studying the probability $\tau_p^{\mathrm{f}}(0, e_n)$, that the origin and the

Function		Behaviour	Exponent
percolation probability	$\theta(p) = P_p(\lvert C\rvert = \infty)$	$\theta(p) \approx (p - p_c)^\beta$	β
truncated mean cluster size	$\chi^f(p) = E_p(\lvert C\rvert; \lvert C\rvert < \infty)$	$\chi^f(p) \approx \lvert p - p_c\rvert^{-\gamma}$	γ
number of clusters per vertex	$\kappa(p) = E_p(\lvert C\rvert^{-1})$	$\kappa'''(p) \approx \lvert p - p_c\rvert^{-1-\alpha}$	α
cluster moments	$\chi_k^f(p) = E_p(\lvert C\rvert^k; \lvert C\rvert < \infty)$	$\dfrac{\chi_{k+1}^f(p)}{\chi_k^f(p)} \approx \lvert p - p_c\rvert^{-\Delta}$	Δ
correlation length	$\xi(p)$	$\xi(p) \approx \lvert p - p_c\rvert^{-\nu}$	ν
cluster volume		$P_{p_c}(\lvert C\rvert = n) \approx n^{-1-1/\delta}$	δ
cluster radius		$P_{p_c}\big(\text{rad}(C) = n\big) \approx n^{-1-1/\rho}$	ρ
connectivity function		$P_{p_c}(0 \leftrightarrow x) \approx \lvert x\rvert^{2-d-\eta}$	η

Table 9.2. Eight functions and their critical exponents.

vertex $e_n = (n, 0, 0, \dots, 0)$ belong to the same finite open cluster. It is believed that

(9.15) $$\tau_p^f(0, e_n) \approx \begin{cases} n^{2-d-\eta} & \text{if } p = p_c, \\ e^{-n/\xi(p)} & \text{if } p \neq 0, p_c, 1, \end{cases}$$

for some $\xi(p)$ satisfying $\xi(p) \to \infty$ as $p \to p_c$. How does $\tau_p^f(0, e_n)$ behave in the double limit as $p \to p_c$ and $n \to \infty$? It is not unreasonable to guess that $\tau_p^f(0, e_n)$ behaves roughly like $n^{2-d-\eta} f(n/\xi(p))$ for some function f when p is near to p_c and n is large; in the light of Theorems (6.44) and (8.53) we may guess that f is (more or less) the negative exponential function, which is to say that

(9.16) $$\frac{\tau_p^f(0, e_n)}{\tau_{p_c}^f(0, e_n)} \approx e^{-n/\xi(p)} \qquad \text{as } p \to p_c \text{ and } n \to \infty,$$

where $\xi(p_c) = \infty$. That is to say, $\tau_p^f(0, e_n)$ and $\tau_{p_c}^f(0, e_n)$ differ significantly only when $n/\xi(p)$ is large. In this sense, $\xi(p)$ is the natural 'length scale' of bond percolation with edge-probability p. For example, suppose that we are told that, for some N, the origin and e_N are in the same finite open cluster of a supercritical percolation process, but we are not told the value of p. If N is small, we have little information about p, whereas if N is large we have reason to deduce that p is near to p_c. How large need N be in order that we may deduce that $p < p_c + \varepsilon$, say? From the above argument, we require N to have at least the order of magnitude of $\xi(p_c + \varepsilon)$. That is to say, as $\varepsilon \downarrow 0$, we require that the order of N grows at least as fast as $\xi(p_c + \varepsilon)$. There is an equally valid argument for subcritical percolation.

Mathematical physicists use the term 'correlation length' rather loosely to mean the 'natural length scale' of the process: the 'correlation length' is the minimal scale on which percolation with edge-probability p differs qualitatively from percolation with edge-probability p_c. In the light of the above discussion, we define the *correlation length* $\xi(p)$ by

$$\xi(p)^{-1} = \lim_{n\to\infty} \left\{ -\frac{1}{n} \log \tau_p^{\mathrm{f}}(0, e_n) \right\}, \tag{9.17}$$

in agreement with the definitions of Sections 6.2 and 8.5. It is a fundamental assumption of scaling theory that there exists a unique length scale in the above sense, and it is neither clear that this holds nor is it self-evident that our definition (9.17) is the most useful definition of correlation length. We note that we are concerned only with the order of magnitude of $\xi(p)$ as p approaches p_c. It would be equally acceptable to define the correlation length to be $2\xi(p)$, and more generally to be any function $\varphi(p)$ with the property that $\varphi(p)/\xi(p)$ is bounded away from 0 and ∞ on a neighbourhood of p_c.

There is another definition of correlation length which has received considerable attention in the literature, and this is the mean radius of a finite open cluster, defined formally by

$$\varphi(p) = \sqrt{\frac{1}{\chi^{\mathrm{f}}(p)} \sum_{x\in\mathbb{Z}^d} |x|^2 \tau_p^{\mathrm{f}}(0, x)} \qquad \text{for } p \neq p_c. \tag{9.18}$$

We may think of $\varphi(p)$ as being $\sqrt{E(|X|^2)}$ where X is a random vertex of $\mathbb{Z}^d$ having distribution

$$P(X = x) = \frac{1}{\chi^{\mathrm{f}}(p)} \tau_p^{\mathrm{f}}(0, x) \qquad \text{for } x \in \mathbb{Z}^d. \tag{9.19}$$

We leave it as an exercise to devise reasonable assumptions under which $\varphi(p)/\xi(p)$ is bounded away from 0 and ∞ on a neighbourhood of p_c.

What is the scaling theory of correlation length? That is to say, what is the asymptotic behaviour of $\xi(p)$ as p approaches p_c from above or beneath? A fundamental assumption of scaling theory is that there exists $\nu > 0$ such that

$$\xi(p) \approx |p - p_c|^{-\nu} \qquad \text{as } p \to p_c, \tag{9.20}$$

and we therefore conjecture that such an asymptotic relation is valid. As before, there are two aspects to this conjecture: that $\xi(p)$ approaches $\xi(p_c) = \infty$ in the manner of a negative power of $|p - p_c|$, and that the power in question does not depend on whether p approaches p_c from above or beneath.

There is a scaling relation involving the correlation length exponent ν:

$$\gamma = \nu(2 - \eta). \tag{9.21}$$

Before moving on to the hyperscaling relations, we note that (9.13), (9.14), and (9.21) provide four equations for seven quantities α, β, γ, δ, Δ, η, ν.

The *hyperscaling relations* involve the number d of dimensions, and assert that

$$d\rho = \delta + 1, \tag{9.22}$$

$$d\nu = 2 - \alpha, \tag{9.23}$$

so that all the terms in (9.13) equal $d\nu$. The scaling and hyperscaling relations combine to form six equations in eight unknowns. It is believed that the hyperscaling relations are valid only for values of d satisfying $d \le d_c$, for some d_c called the *(upper) critical dimension*. For percolation on $\mathbb{L}^d$ when $d \ge d_c$, it is believed that the process behaves roughly in the same manner as percolation on an infinite regular tree (that is, a lattice without circuits). More precisely, many believe that the critical exponents take on the corresponding values from percolation on a regular tree when $d \ge d_c$; we shall see in Section 10.1 that

$$\alpha = -1,\ \beta = 1,\ \gamma = 1,\ \delta = 2,\ \Delta = 2,\ \rho = \tfrac{1}{2},\ \eta = 0,\ \nu = \tfrac{1}{2},$$

for such a percolation process. If these values are attained by percolation on $\mathbb{L}^d$ and the hyperscaling relation $d\nu = 2 - \alpha$ is valid, then $d = 6$. This calculation provides some evidence that the hyperscaling relations are valid only when $d \le 6$ and that, in six or more dimensions, the critical exponents take on the values corresponding to percolation on a regular tree. Section 10.3 contains a summary of progress towards the latter statement for large d.

The critical exponents α, β, γ, δ, Δ, η, ρ, ν depend of course upon the number d of dimensions, but there are physical reasons to suppose that they do not depend on the particular lattice structure. That is to say, it is believed that the critical exponents of bond or site percolation on any d-dimensional lattice $\mathcal{L}$ are the same as those for bond percolation on $\mathbb{L}^d$. This belief has much of its basis in the hypothesis that only over large 'length scales' do critical phenomena manifest themselves, and the local lattice structure is not easily observed over such distances. There is considerable numerical evidence to support this conjecture of 'universality', which may be phrased as follows. Consider the class of all percolation processes on d-dimensional lattices having finite vertex degrees. It is believed that this is a 'universality class' of processes, in the sense that the natures of the phase transitions thereof are similar; in particular, all processes in this class are believed to have equal critical exponents.

We may summarize the ideas of this section as follows.

(a) Macroscopic quantities such as θ, χ^{f}, and κ''' behave in the manner of powers of $|p - p_{\mathrm{c}}|$ when p is near to p_{c}.

(b) The distributions of the radius and size of open clusters have tails which decay as negative powers when $p = p_{\mathrm{c}}$.

(c) There is a fundamental minimal scale $\xi(p)$ over which percolation with edge-probability p is distinguishable from percolation with $p = p_{\mathrm{c}}$, and this 'correlation length' behaves like a negative power of $|p - p_{\mathrm{c}}|$ when p is near to p_{c}.

(d) The critical exponents satisfy the scaling relations.

(e) The critical exponents satisfy the hyperscaling relations if $d \le 6$ and take on the values corresponding to percolation on a regular tree when $d \ge 6$.

(f) The critical exponents are 'universal' in the sense that they depend only on the number d of dimensions and not otherwise upon the individual structure of the underlying lattice.

In the next section, we present a more detailed account of scaling theory, leading to derivations of the scaling and hyperscaling relations. Underlying scaling theory is the theory of renormalization, and Section 9.3 is a superficial essay on this topic.

The asymptotic analysis of this chapter will appear rather shaky to some, and not without good reason. We shall use three types of notation to denote asymptotic equivalence. We write $a_n \approx b_n$ (respectively $a(p) \approx b(p)$) if $\log a_n / \log b_n \to 1$ as $n \to \infty$ (respectively $\log a(p)/\log b(p) \to 1$ as $p \to p_{\mathrm{c}}$). We write $a_n \sim b_n$ if $a_n/b_n \to 1$ as $n \to \infty$, and similarly for functions of p as $p \to p_{\mathrm{c}}$. The notation $\simeq$ is reserved for use in a mathematically imprecise manner to relate pairs of quantities whose asymptotic behaviours are believed to be rather similar.

9.2 Scaling Theory

The principal hypothesis of scaling theory is that quantities such as $P_p(|C| = n)$ have certain specific asymptotic expansions which are valid for all p near to p_{c} and for all large n. Starting from such hypotheses, we may use scaling theory to arrive at the scaling relations (9.13), (9.14), and (9.21). Other hypotheses yield the hyperscaling relations. We describe such an approach in this section; in the next section, we shall describe briefly how the theory of renormalization may be used to provide a heuristic basis for scaling theory.

We assume henceforth that all the critical exponents $\alpha, \beta, \gamma, \delta, \Delta, \eta, \rho, \nu$ are defined as in the last section. Our first scaling hypothesis is that

$$(9.24)\qquad P_p(|C| = n) \sim \begin{cases} n^{-\sigma} f_-(n/\xi(p)^\tau) & \text{if } p \le p_{\mathrm{c}}, \\ n^{-\sigma} f_+(n/\xi(p)^\tau) & \text{if } p \ge p_{\mathrm{c}}, \end{cases}$$

where σ and τ are positive constants, f_- and f_+ are smooth functions (infinitely differentiable, say) on $[0, \infty)$, and the relation $\sim$ is understood in the usual asymptotic sense to mean that the ratio of the left side to the right side converges to 1 as $n \to \infty$ and $p \to p_c$; note that the notation σ, τ is not standard. Following Theorems (6.78), (8.61), and (8.65), it is not unreasonable to guess that $f_-(m) \approx \exp(-Am)$ and $f_+(m) \approx \exp(-Bm^{(d-1)/d})$ for some A and B as $m \to \infty$, but we shall not assume this here. Instead, we shall suppose only that $f_-(0) = f_+(0) > 0$, and that $f_-(m) \to 0$ and $f_+(m) \to 0$ faster than any power of m^{-1} as $m \to \infty$. The first of these assumptions guarantees that the two asymptotic relations in (9.24) coincide when $p = p_c$, and the second implies that the behaviour of $P_p(|C| = n)$ is dominated when $p \neq p_c$ by a quantity which depends only on the ratio of n to a power of the correlation length.

In the matter of terminology, we note that physicists commonly refer to such a hypothesis as (9.24) as being an 'ansatz'. There are certain differences between an ansatz and a hypothesis, the former being often a guessed formula which may be used as a basis for calculations.

The asymptotic analysis which follows is somewhat speculative. Consequently, the correct form of scaling hypothesis is not entirely obvious. It is not uncommon to encounter in the literature a version of (9.24) which is assumed to be valid in the limit when $n \to \infty$ and $p \to p_c$ in such a way that $n/\xi(p)^\tau$ approaches a positive constant.

This scaling hypothesis (or ansatz) introduces the new exponents σ and τ, but these are readily expressed in terms of those exponents already to hand. We remember that $P_{p_c}(|C| = n) \approx n^{-1-1/\delta}$, and we deduce that

$$\sigma = 1 + \delta^{-1}. \tag{9.25}$$

We next cross our fingers and perform the following calculations, typical of scaling theory. In the following, f represents f_- if $p < p_c$ and f_+ if $p > p_c$. We have that

$$\begin{aligned} \chi^{\mathrm{f}}(p) &= \sum_n n P_p(|C| = n) \simeq \sum_n n^{1-\sigma} f(n\xi(p)^{-\tau}) \\ &\simeq \int_0^\infty n^{1-\sigma} f(n\xi(p)^{-\tau})\, dn \\ &= \xi(p)^{\tau(1-\sigma)+\tau} \int_0^\infty u^{1-\sigma} f(u)\, du \end{aligned} \tag{9.26}$$

by the substitution $u = n\xi(p)^{-\tau}$. Now $\sigma = 1 + \delta^{-1}$ and $\delta \geq 1$, giving that $1 - \sigma = -\delta^{-1} \geq -1$. If $\delta > 1$, then $1 - \sigma > -1$ and the last integral converges, but we have not eliminated the case $\delta = 1$, for which $1 - \sigma = -1$ and the integral diverges. We shall see in Section 10.2 that $\delta \geq 2$ whenever δ exists, so that this discussion is academic; an alternative argument uses the assumption that $\chi^{\mathrm{f}}(p) < \infty$ when $p \neq p_c$. We now make the assumptions that $\chi^{\mathrm{f}}(p) \approx |p - p_c|^{-\gamma}$ and $\xi(p) \approx |p - p_c|^{-\nu}$ to find that

$$\gamma = \nu\tau(2 - \sigma)$$

which, in conjunction with (9.25), yields

$$\tau = \frac{\gamma}{\nu(1-\delta^{-1})}, \tag{9.27}$$

and we have expressed σ and τ in terms of γ, δ, ν.

We move on to consider the percolation probability $\theta(p)$, and we suppose now that $p > p_c$. Then

$$\begin{aligned} \theta(p) &= 1 - \sum_n P_p(|C| = n) \\ &= \sum_n \{P_{p_c}(|C| = n) - P_p(|C| = n)\} \qquad \text{if } \theta(p_c) = 0 \\ &\simeq \sum_n n^{-\sigma}\{f_+(0) - f_+(n\xi(p)^{-\tau})\} \\ &\simeq \int_0^\infty n^{-\sigma}\{f_+(0) - f_+(n\xi(p)^{-\tau})\}\,dn \\ &= \xi(p)^{\tau(1-\sigma)} \int_0^\infty u^{-\sigma}\{f_+(0) - f_+(u)\}\,du \end{aligned} \tag{9.28}$$

as before; such asymptotic relations should not be interpreted too strictly. The last integral converges since the integrand behaves like $-u^{1-\sigma} f_+'(0)$ near $u = 0$. From the assumption $\theta(p) \approx (p - p_c)^\beta$, we have that

$$\beta = -\nu\tau(1-\sigma),$$

giving from (9.25) and (9.27) that

$$\gamma + 2\beta = \beta(\delta + 1), \tag{9.29}$$

our first scaling relation.

Next we apply the same argument to the number $\kappa(p)$ of open clusters per vertex. As before, f represents f_- if $p < p_c$ and f_+ if $p > p_c$. We have that

$$\begin{aligned} \kappa(p) &= \sum_n \frac{1}{n} P_p(|C| = n) \simeq \sum_{n=1}^{\infty} n^{-1-\sigma} f(n\xi(p)^{-\tau}) \\ &\simeq \xi(p)^{-\tau\sigma} \int_{\xi(p)^{-\tau}}^\infty u^{-1-\sigma} f(u)\,du \end{aligned} \tag{9.30}$$

as before. Now, if ε is small and positive,

(9.31)

$$\begin{aligned} \int_{\xi(p)^{-\tau}}^\infty u^{-1-\sigma} f(u)\,du &= \int_{\xi(p)^{-\tau}}^{\varepsilon} u^{-1-\sigma} f(u)\,du + \int_\varepsilon^\infty u^{-1-\sigma} f(u)\,du \\ &\simeq f(0)\left[-\frac{1}{\sigma} u^{-\sigma}\right]_{\xi(p)^{-\tau}}^{\varepsilon} + \int_\varepsilon^\infty u^{-1-\sigma} f(u)\,du \\ &= A\xi(p)^{\tau\sigma} + B, \end{aligned}$$

where A and B are constants which depend upon ε. We substitute this into (9.30) to find that

$$\kappa(p) \simeq A + B\xi(p)^{-\tau\sigma} \simeq A + B|p - p_c|^{\nu\tau\sigma}. \tag{9.32}$$

We compare this with the assumption $\kappa'''(p) \approx |p - p_c|^{-1-\alpha}$, and we perform a sufficiently large leap of the imagination to arrive at

$$-1 - \alpha = \nu\tau\sigma - 3.$$

We now use (9.25), (9.27), and (9.29) to deduce the second scaling relation:

$$2 - \alpha = \gamma + 2\beta. \tag{9.33}$$

We argue similarly to obtain

$$E_p(|C|^k; |C| < \infty) \simeq D_k\xi(p)^{\tau(k+1-\sigma)}$$

if $k \geq 1$, where D_k is a constant which depends on k. Comparison with (9.7) yields a value for the gap exponent:

$$\begin{aligned} \Delta &= \tau\nu \\ &= \delta\beta \qquad \text{by (9.27) and (9.29).} \end{aligned} \tag{9.34}$$

This is just about as far as we may progress from the scaling hypothesis (9.24) for the cluster size distribution. The easiest way to obtain the relation (9.21), $\gamma = \nu(2 - \eta)$, is to introduce a new scaling ansatz; we hypothesize that

$$\tau_p^{\mathrm{f}}(0, x) \simeq \begin{cases} |x|^{2-d-\eta} g_-(|x|/\xi(p)) & \text{as } p \uparrow p_c \text{ and } |x| \to \infty, \\ |x|^{2-d-\eta} g_+(|x|/\xi(p)) & \text{as } p \downarrow p_c \text{ and } |x| \to \infty, \end{cases} \tag{9.35}$$

for some smooth functions g_- and g_+ (infinitely differentiable, say) satisfying $g_-(0) = g_+(0) > 0$ and $g_-(m) \to 0$, $g_+(m) \to 0$ faster than any power of m^{-1} as $m \to \infty$. Taking (9.15) into account, it is reasonable to suppose that $g_-(m) \approx e^{-Am}$ and $g_+(m) \approx e^{-Bm}$ as $m \to \infty$ for constants A and B, but we shall not assume this here. We have used the undefined asymptotic relation $\simeq$ in (9.35) for the following reason. Presumably $\tau_p^{\mathrm{f}}(0, x)$ depends not only on the distance $|x|$ from the origin to x, but also on the geometry of the position of x. However, it is likely that we may re-scale $|x|$ in an appropriate way depending on this geometry, with the effect that equation (9.35), or something very similar, becomes valid with $\simeq$ replaced by $\sim$ and $|x|$ replaced by the re-scaled distance function; we do not attempt this here.

Summing (9.35) over all x, we obtain

$$\chi^{\mathrm{f}}(p) \simeq \sum_x |x|^{2-d-\eta} g(|x|/\xi(p)),$$

where g is interpreted as g_- or g_+ depending on whether $p < p_c$ or $p > p_c$. The number of vertices x for which $|x| = n$ has order n^{d-1}, so that

$$\begin{aligned}\chi^f(p) &\simeq A \sum_n n^{d-1} n^{2-d-\eta} g(n/\xi(p)) \\ &\simeq A \int_0^\infty n^{1-\eta} g(n/\xi(p))\, dn \\ &\simeq A\xi(p)^{2-\eta} \int_0^\infty u^{1-\eta} g(u)\, du\end{aligned}$$

for some constant A. It cannot be the case that $\eta \geq 2$, since $\chi^f(p)$ and $\xi(p)$ are assumed to diverge at $p = p_c$ in the manner of negative powers of $|p - p_c|$. Thus $\eta < 2$ and the integral converges, implying that

$$\chi^f(p) \simeq B\xi(p)^{2-\eta}$$

for some constant B; we set the exponents of $|p - p_c|$ to be equal, and obtain

$$\gamma = \nu(2 - \eta), \tag{9.36}$$

our last scaling relation.

We turn now to the hyperscaling relations, beginning with the relation (9.22), $d\rho = \delta + 1$, involving the critical exponent ρ for the cluster radius. Let $p = p_c$, and let $\mathrm{rad}(C) = \max\{|x| : x \in C\}$ be the radius of the open cluster at the origin. If $\mathrm{rad}(C) \geq k$ say, then how large is the volume $|C|$? The following is a rather crude approach to this question. Suppose that $\mathrm{rad}(C) \geq k$ and let us consider vertices x of $\mathbb{L}^d$ satisfying $|x| \leq k$. It may not be too unreasonable to guess that the proportion of such vertices x which are joined to the origin is roughly equal to the probability that such a vertex belongs to an open cluster having radius at least k; this probability is approximately $k^{-1/\rho}$, so that

$$|C| \geq k^d k^{-1/\rho} \quad \text{if and only if} \quad \mathrm{rad}(C) \geq k, \tag{9.37}$$

to within a few orders of magnitude and some error probability. Now,

$$\begin{aligned}P_{p_c}\big(n \leq |C| < \infty\big) &\approx n^{-1/\delta}, \\ P_{p_c}\big(k \leq \mathrm{rad}(C) < \infty\big) &\approx k^{-1/\rho},\end{aligned}$$

so that (9.37) implies

$$\left(d - \frac{1}{\rho}\right)\frac{1}{\delta} = \frac{1}{\rho},$$

or

$$d\rho = \delta + 1 \tag{9.38}$$

as required.

The following argument yields the second hyperscaling relation (9.23), which asserts that $d\nu = 2 - \alpha$. Let $T(n) = [1, n]^d$. In the calculations (9.30)–(9.32) we saw that the 'singular' part of $\kappa(p)$ is the contribution from open clusters with sizes having orders of $\xi(p)^\tau$ or greater. Let $K_n(\xi(p)^\tau)$ be the number of such open clusters in the box $T(n)$; this is not a particularly rigorous definition, but its style is in keeping with the rest of the argument. If we set $n = \xi(p)$, the correct 'length scale' for percolation with edge-probability p, then the mean number $E_p\{K_n(\xi(p)^\tau)\}$ of such clusters should not vary greatly for values of p near to p_c. Suppose then that

$$E_p\big\{K_{\xi(p)}(\xi(p)^\tau)\big\} \simeq A \qquad \text{as } p \to p_c,$$

for some constant A. For any positive integer m, we may think of $T(m\xi(p))$ as comprising m^d copies of $T(\xi(p))$. Only few of the open clusters of $T(m\xi(p))$ with size $\xi(p)^\tau$ or greater intersect two or more of these copies, so that

$$E_p\big\{K_{m\xi(p)}(\xi(p)^\tau)\big\} \simeq m^d E_p\big\{K_{\xi(p)}(\xi(p)^\tau)\big\} \simeq Am^d. \tag{9.39}$$

The 'singular' part of $\kappa(p)$ is given therefore by

$$\lim_{m\to\infty}\left\{\frac{1}{(m\xi(p))^d} E_p\big\{K_{m\xi(p)}(\xi(p)^\tau)\big\}\right\} \simeq A\xi(p)^{-d}$$

by (9.39). We deduce by another leap of the imagination that

$$2 - \alpha = d\nu, \tag{9.40}$$

the controversial hyperscaling relation.

We summarize the discussion of this section as follows.

(a) There are two attractive scaling hypotheses concerning the behaviour of $P_p(|C| = n)$ and $\tau_p^{\mathrm{f}}(0, x)$ when n and $|x|$ are large, and p is near to p_c. These scaling hypotheses provide strong evidence for the scaling relations.

(b) Two further heuristic arguments lead to the hyperscaling relations.

9.3 Renormalization

Renormalization techniques are of great value in statistical mechanics and have been applied with some success to percolation models. They provide a unified basis for scaling theory which is attractive in its simplicity and coherence. However, with the exception of a handful of special cases, we know of no completely rigorous and major application of renormalization theory which contributes to an understanding of scaling theory for percolation in two or more dimensions. This

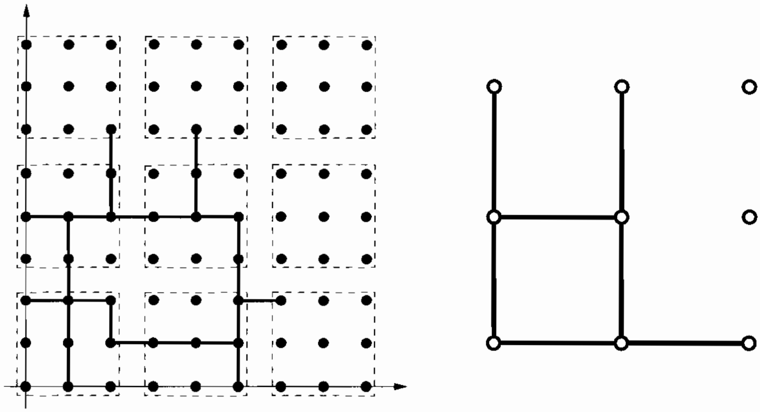


Figure 9.1. Renormalizing the square lattice with $b = 3$. The lines in the left picture represent open edges, whilst the right picture contains the 'renormalized' vertices together with edges which describe the traversibility of the 3×3 blocks of the original lattice.

notwithstanding, the intuitive appeal of 'real-space renormalization' is considerable and, for many people, such arguments are the cornerstone of the theory of critical phenomena. We shall attempt to sketch this approach.

Before continuing, we note that 'block' arguments have proved very useful in understanding percolation when $p \neq p_c$. For example, such an argument was used for the proof in Section 7.2 that p_c equals the limit of the critical probabilities of slabs. Block arguments employ a type of renormalization, but they have not yet contributed substantially to a rigorous theory of the critical phenomenon.

Here is an example of a renormalization argument. We consider bond percolation on $\mathbb{L}^d$ where $d \geq 2$, and we let b be a positive integer satisfying $b \geq 2$. We partition $\mathbb{Z}^d$ into cubes of size b^d in a natural way, the sets of this partition having the form

$$B(j_1, j_2, \dots, j_d) = \prod_{i=1}^{d} \big[j_i b, (j_i + 1)b - 1\big]$$

as $j = (j_1, j_2, \dots, j_d)$ ranges over the set of all d-vectors of integers; see Figure 9.1 for a two-dimensional sketch. We call the $B(j)$ *blocks*, and we say that the block $B(j_1, \dots, j_d)$ is *traversible in the rth direction* if there exists an open path in the block joining some vertex x with $x_r = j_r b$ to some vertex y with $y_r = (j_r + 1)b - 1$, and also the edge leading out of y in the direction of the positive rth coordinate axis is open.

We now 'renormalize' the lattice, by replacing each block $B(j_1, \dots, j_d)$ by a single vertex which we label $j = (j_1, \dots, j_d)$. In this new lattice, we place edges between all pairs j, k of vertices for which

$$\sum_i |j_i - k_i| = 1,$$

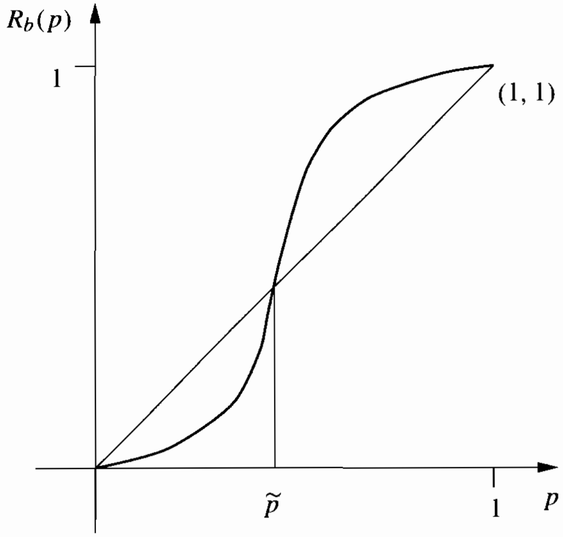


Figure 9.2. A sketch of the function R_b with emphasis on its non-trivial fixed point $\widetilde{p}$. Note the characteristic 'S-shape' of the function, and the fact that the derivative $\lambda = R_b'(\widetilde{p})$ at $\widetilde{p}$ satisfies $\lambda > 1$.

so that the new lattice is a copy of $\mathbb{L}^d$; we denote the new lattice by $\mathcal{R}_b(\mathbb{L}^d)$. We declare the edges of $\mathcal{R}_b(\mathbb{L}^d)$ to be open or closed according to the following rule: the edge joining the two new vertices $(j_1, \dots, j_{r-1}, j_r, j_{r+1}, \dots, j_d)$ and $(j_1, \dots, j_{r-1}, j_r + 1, j_{r+1}, \dots, j_d)$ is declared to be open if and only if the block $B(j_1, \dots, j_d)$ is traversible in the rth direction. Thus each edge of $\mathcal{R}_b(\mathbb{L}^d)$ is open with probability $p^* = R_b(p)$, where $R_b(p)$ is the probability that $B(0, \dots, 0)$ is traversible in the first direction. We illustrate this construction in Figure 9.1, and we sketch the function R_b in Figure 9.2.

In the original lattice, the states of different edges were independent of each other. This is not true in the new lattice, since the state of each edge contains information about the states of certain other edges having a vertex in common. *Let us suppose however that the new process resembles closely a realization of bond percolation with edge-probability* p^*. If this were the case, the process of renormalization would replace a bond percolation process having edge-probability p by a bond percolation process on a re-scaled lattice having edge-probability $p^* = R_b(p)$. *Let us suppose next that the large-scale connectivities of these two processes are very similar.*

Whither are we led by these two assumptions? We consider first the correlation length $\xi(p)$ of the first lattice; this is the fundamental 'length scale' of the first process, being the shortest scale over which the process can be distinguished from a critical percolation process with edge-probability p_c. In the new process, the correlation length is $\xi(p^*)$. However, each unit of length of the new lattice $\mathcal{R}_b(\mathbb{L}^d)$

corresponds to b units of length of the first, and we have supposed that the two processes are similar over long ranges. This suggests the equation

$$\xi(p) = b\xi(p^*) \tag{9.41}$$

where

$$p^* = R_b(p). \tag{9.42}$$

We now cross our fingers, and let $\widetilde{p}$ be a fixed point of the mapping R_b. If $\widetilde{p} = R_b(\widetilde{p})$ then $\xi(\widetilde{p}) = b\xi(\widetilde{p})$, so that $\xi(\widetilde{p})$ equals 0 or ∞. However, $\xi(\widetilde{p}) = 0$ if and only if $\widetilde{p}$ equals 0 or 1, and these fixed points are of little interest. It follows that the remaining fixed point of R_b is the critical probability p_c, since this is the unique value of $\widetilde{p}$ for which $\xi(\widetilde{p}) = \infty$. We deduce therefore that p_c is a root of the equation $R_b(p) = p$. When $b = 2$ say, the function R_b is easily calculated in closed form. Needless to say, it does not turn out in practice that $p_c = R_b(p_c)$, and the reason for this is that the two italicized hypotheses above can be considered to be rough guidelines only about the behaviour of the new process on the renormalized lattice. The process on the lattice $\mathcal{R}_b(\mathbb{L}^d)$ is *not* bond percolation, and the large-scale properties of the two lattices will generally differ. There are at least two directions which may be pursued from this point. The first is to guess that, as b grows, the fixed point of R_b approximates better and better to p_c; this argument has met with some numerical success, although current technology allows the estimation of R_b by Monte Carlo techniques only, when b is large. The second direction is theoretical rather than practical, and this is to attempt to describe and analyse with greater care the random process on the renormalized lattice. We do not pursue this here, but refer the reader to Fisher (1983) and Kesten (1987e, Section 2.3).

We return to the equations (9.41) and (9.42) which describe the effect of renormalization upon the correlation length $\xi(p)$, and we suppose that p_c is the unique non-trivial fixed point of the function R_b. The behaviour of $\xi(p)$ when p is near to p_c is related to the corresponding behaviour of $R_b(p)$ near to its fixed point p_c, and so we note that

$$R_b(p) - R_b(p_c) = \lambda(p - p_c) + o(|p - p_c|) \tag{9.43}$$

where $\lambda = R_b'(p_c)$. Now, $R_b(p)$ is the probability of an increasing event which depends on only finitely many edges. We may therefore apply Theorem (2.36), and specifically inequality (2.37), to deduce that

$$\lambda = R_b'(p_c) \geq \frac{R_b(p_c)(1 - R_b(p_c))}{p_c(1 - p_c)} = 1.$$

It is not difficult to see from the proof of (2.41) that strict inequality holds here, which is to say that $\lambda > 1$. We express (9.43) in the form

$$p^* - p_c \sim \lambda(p - p_c) \qquad \text{as } p \to p_c \tag{9.44}$$

to see that the process of renormalization pushes the percolation process away from the critical probability, in the sense that

$$|p^* - p_c| \geq \lambda\big(1 + o(1)\big)|p - p_c|$$

when $|p - p_c|$ is small. We substitute (9.44) into (9.41) to find that $\psi(p - p_c) = \xi(p)$ satisfies

(9.45) $$\psi(p - p_c) \simeq b\psi\big(\lambda(p - p_c)\big) \qquad \text{when } |p - p_c| \text{ is small.}$$

Let A be a positive constant and suppose that $|p - p_c|$ is very small compared with A. We iterate (9.45) m times to obtain

(9.46) $$\psi(p - p_c) \simeq b^m \psi\big(\lambda^m (p - p_c)\big)$$

and we choose m such that $\lambda^m |p - p_c| \simeq A$. With this value of m, (9.46) becomes

(9.47) $$\begin{aligned} \psi(p - p_c) &\simeq \psi(A) \exp\left(\frac{\log b}{\log \lambda}\big(\log A - \log|p - p_c|\big)\right) \\ &= D_1 |p - p_c|^{-\nu}, \end{aligned}$$

where D_1 is a constant which depends on b, λ, and A, and

(9.48) $$\nu = \frac{\log b}{\log \lambda}.$$

We have shown that the equation $\xi(p) = b\xi(p^*)$, where $p^* = R_b(p)$, implies that $\xi(p)$ behaves in the manner of a power of $|p - p_c|$ when p is near to p_c, and furthermore it specifies the numerical value of the exponent ν. This 'exact calculation' of ν is only an approximation to the actual value of ν, but the error is expected to be small in the limit as $b \to \infty$.

We have based the argument so far upon the effect of renormalization upon the correlation length. Parallel analyses are possible for such quantities as the connectivity functions and the number of open clusters per vertex. For example, if we concentrate on the latter function $\kappa(p)$, we obtain

(9.49) $$\kappa(p) \simeq b^{-d} \kappa(p^*)$$

in place of $\xi(p) = b\xi(p^*)$, and this relation yields

(9.50) $$\kappa(p) \simeq D_2 |p - p_c|^{\nu d},$$

where D_2 is a constant and ν is given by (9.48). As before, we have deduced that κ behaves roughly in the manner of a power of $|p - p_c|$ when p is near to p_c; combined with the hypothesis (9.6) that $\kappa'''(p) \approx |p - p_c|^{-1-\alpha}$, we may deduce that $2 - \alpha = d\nu$, in agreement with the hyperscaling relation (9.23).

We may use the 'exact value' (9.48) of ν to find the 'exact value' of α. Care is advisable in interpreting the asymptotic relation in (9.50); for example, contrary to appearances, it is not the case that $\kappa(p_c) = 0$.

The theory of renormalization lends support to the hypothesis of universality. We may think of renormalization as an operator acting on a class of processes including d-dimensional percolation models. The values of critical exponents should depend only on the re-scaling factor b and the behaviour of this operator near to its fixed point—see (9.48), for example. These values should thus depend upon the number d of dimensions but not upon other features of individual processes within a large class of 'similar' processes.

It is not difficult to pursue such arguments further, and we refer the reader to Essam (1980), Stauffer (1979), Stauffer and Aharony (1991), Fisher (1983), Reynolds, Stanley, and Klein (1980), Stanley, Reynolds, Redner, and Family (1982), Hughes (1996), and the references therein.

9.4 The Incipient Infinite Cluster

The term 'incipient infinite cluster' has been used in the physics literature in an attempt to describe the infinite open cluster when $p = p_c$. Unfortunately, no such cluster exists when $d = 2$ or $d \geq 19$, and it is widely believed that this is the case whenever $d \geq 2$. How then may we make mathematical sense of the physical idea of the infinite open cluster at the critical point p_c? Kesten (1986b) has made the following proposal in the case of two dimensions, and parallel developments may be valid in higher dimensions. Suppose that $d = 2$. The event $\{|C| = \infty\}$, that the origin lies in an infinite open cluster, has probability 0 when $p = p_c$. We may however force this event to occur by suitable conditioning, and this may be done in two ways.

First we take $p > p_c$, and we define the probability measure ν_p by

(9.51) $$\nu_p(A) = P_p\big(A \,\big|\, |C| = \infty\big)$$

for cylinder events A. Kesten (1986b) has proved that the limit of $\nu_p(A)$ exists as $p \downarrow p_c$; denoting the limit by $\nu(A)$, we have that

(9.52) $$\nu(A) = \lim_{p \downarrow p_c} \nu_p(A)$$

for cylinder events A.

Secondly, we take $p = p_c$ and let $B(n)$ be the usual box having side-length $2n$ and centre at the origin. For cylinder events A, we write

(9.53) $$\nu_n(A) = P_{p_c}\big(A \,\big|\, 0 \leftrightarrow \partial B(n)\big)$$

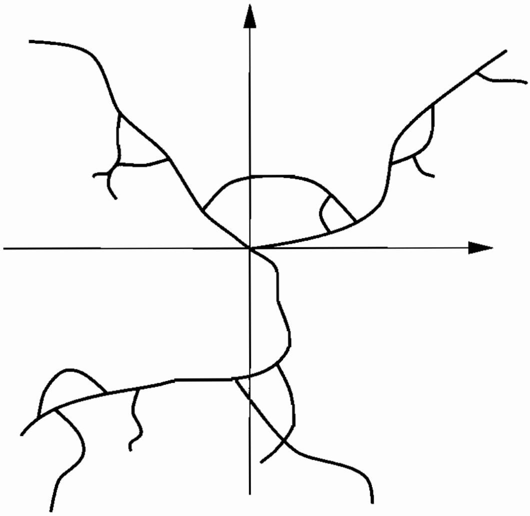


Figure 9.3. A sketch of a sparse open cluster having connections over long ranges but few interconnections between distant parts.

for the probability of A conditional on there being an open path from the origin to the surface of $B(n)$. It turns out that the limit of $\nu_n(A)$ exists as $n \to \infty$, and that this limit coincides with the limit in (9.52). That is to say,

$$\lim_{n\to\infty} \nu_n(A) = \nu(A) \tag{9.54}$$

for cylinder events A. The limit function ν is defined on cylinder events only, but it has a unique extension to the σ-field $\mathcal{F}$.

It is clear by the nature of the conditioning in (9.51) and (9.53) that the origin has ν-probability 1 of belonging to an infinite open cluster I, and it turns out that this cluster is (ν-a.s.) the unique infinite open cluster of the lattice. We call this infinite cluster the *incipient infinite cluster*, although this expression refers in reality to the *distribution* of I under ν, rather than to I itself.

It is an open problem both to prove that $\theta(p_c) = 0$ and to give a rigorous justification of the above argument in three or more dimensions under the hypothesis that $\theta(p_c) = 0$.

Physicists are interested in the geometry of the incipient infinite cluster, and particularly in its properties of statistical self-similarity over different length-scales. Little rigorous progress has been made in this direction, even in the case of two dimensions. Presumably the incipient infinite cluster permeates the space which contains it, but is rather sparse within this space. It is believed to have zero density but to be extremely 'ramified', which is to say that it contains long dendritic connections across space but relatively few interconnections; see Figure 9.3 for a sketch of such a network in two dimensions.

A first step towards understanding the geometry of I in two dimensions is to estimate the density of I. Towards this end, we ask for an estimate of the probability that a vertex x of $\mathbb{Z}^2$ lies in I. Let m be a positive integer and suppose that x is a vertex with $\|x\| = m$. It may not be unreasonable to guess that, for large n and 'most' such vertices x,

$$\nu_n(0 \leftrightarrow x) = P_{p_c}\big(0 \leftrightarrow x \,\big|\, 0 \leftrightarrow \partial B(n)\big) \simeq m^{-1/\rho},$$

since, conditional on there existing an open path from the origin to a vertex of $\partial B(n)$, the vertex x is required to be connected to vertices in the open cluster at the origin, and some of these vertices are approximately distance m from x. We let n tend to infinity to obtain

(9.55) $$\nu(x \in I) \simeq \|x\|^{-1/\rho} \qquad \text{for } x \in \mathbb{Z}^2.$$

The number of vertices of $B(n)$ which lie in I is approximately

(9.56) $$\begin{aligned} E_\nu\big(|B(n) \cap I|\big) &= \sum_{x \in B(n)} \nu(x \in I) \\ &\simeq \sum_{x \in B(n)} \|x\|^{-1/\rho} \\ &\simeq D_1 \sum_{m=1}^{n} m \cdot m^{-1/\rho} \\ &\simeq D_2 n^{2-1/\rho} \end{aligned}$$

for some constants D_1 and D_2, where E_ν denotes the expectation operator associated with the measure ν. This rough argument indicates that the density of I in $B(n)$ behaves in the manner of $n^{-1/\rho}$ as $n \to \infty$, and furthermore we may think of I as being a subset of $\mathbb{L}^2$ having dimension $2 - 1/\rho$; this last number is sometimes referred to as the 'fractal dimension' of the incipient infinite cluster. Kesten (1986b) has shown how to make this argument rigorous in two dimensions, in as much as this is possible without assuming the existence of the critical exponent ρ. The corresponding argument for d dimensions indicates a value of $d - 1/\rho$ for the hypothetical 'fractal dimension' of the hypothetical incipient infinite cluster.

It has been suggested by de Gennes (1976) that the asymptotic properties of a random walk on the incipient infinite cluster I may give insight into the geometry of I and particularly its conductivity when it is viewed as an infinite random electrical network. Thinking of the random walker as an 'ant in a labyrinth', we may be interested in such quantities as the mean square displacement of the ant after it has taken n steps. Let X_n be the position of such an ant after n steps, and suppose that $X_0 = 0$. Kesten (1986c) has shown that the random walk is *subdiffusive* in two dimensions, in the sense that

(9.57) $$E(X_n^2) \le A n^{1-\varepsilon}$$

for some constants A and $\varepsilon > 0$, and for all n; we recall that symmetric random walk on $\mathbb{Z}^2$ has mean square displacement equal to n, so that (9.57) asserts that the ant diffuses at a rate which is of a lesser order of magnitude than that of an unrestricted random walk. To see the intuitive reasoning behind this, we introduce the idea of the *backbone* $\mathcal{B}$ of I, being the set of vertices x of I from which there exist both an infinite open path and an open path to the origin which is edge-disjoint from the first path (note that this definition of the backbone is somewhat non-standard, other authors defining the backbone as the set of vertices from which there exist two or more edge-disjoint infinite open paths). The idea is that the ant divides its time between occupying vertices in the backbone $\mathcal{B}$ and occupying other vertices. However, time spent outside $\mathcal{B}$ is wasted since the ant must return to the backbone in order to progress substantially in its march about the lattice (see Figure 9.3). It turns out that the number of vertices of $B(n)$ which belong to the backbone is approximately $n^2 P_{p_c}(A_n)$ where A_n is the event that there exist two or more edge-disjoint open paths from the origin to the surface of $B(n)$. We have from the BK inequality that

$$P_{p_c}(A_n) \le P_{p_c}\big(0 \leftrightarrow \partial B(n)\big)^2 \approx n^{-2/\rho},$$

if ρ exists. Thus the volume of $\mathcal{B} \cap B(n)$ has at most order $n^{2-2/\rho}$, which is smaller than the order $n^{2-1/\rho}$ of $I \cap B(n)$. This difference in orders of magnitude is the principal reason for the subdiffusive behaviour (9.57) of the ant. See Kesten (1986b, c) for more details.

9.5 Notes

Sections 9.1 and 9.2. The principal reviews of scaling theory are those of Essam (1980), Stauffer (1979), Stauffer and Aharony (1991), and Hughes (1996). We shall not attempt an accurate historical bibliography here, but refer the reader to the references contained in these extensive reviews. We defer a discussion of rigorous results until Chapters 10 and 11, but note that Kesten (1987b) has proved the validity of all scaling relations in two dimensions which do not involve α, subject to the assumption that ρ and ν exist. In related work, Kesten (1987c) has proved the hyperscaling relation $d\rho = \delta + 1$ in the case $d = 2$, subject to a similar assumption of existence. Further work on hyperscaling relations may be found in Borgs, Chayes, Kesten, and Spencer (1997).

The reviews of Aizenman (1987), Kesten (1987e), and Newman (1987a) contain interesting discussions concerning critical exponents and scaling theory in the context of percolation.

The predictions of theoretical physics do not stop at the scaling and hyperscaling relations. Certain exact values are conjectured for critical exponents in two dimensions: $\alpha = -\frac{2}{3}$, $\beta = \frac{5}{36}$, $\gamma = \frac{43}{18}$, $\delta = \frac{91}{5}$, and so on. See the notes for Section 10.2.

It was Toulouse (1974) who discovered the argument that the (upper) critical dimension equals 6. For later work concerning the critical dimension, see Tasaki (1987a, b, 1989), and Chayes and Chayes (1987b). Berlyand and Wehr (1995, 1997) have investigated further the definitions of correlation length.

For a general introduction to the theory of phase transitions, we refer the reader to the book by Stanley (1971).

Section 9.3. Fisher (1983) has written a beautiful account of renormalization, together with a discussion of the hypothesis of universality and the failure of the hyperscaling laws in large dimensions. See also Kesten (1987e) and the references listed at the end of Section 9.3.

Section 9.4. Other approaches to the 'incipient infinite cluster' are contained in Chayes and Chayes (1986a), Chayes, Chayes, and Durrett (1987), Coniglio (1985), Ben-Avraham and Havlin (1982), Kapitulnik, Aharony, Deutscher, and Stauffer (1983), Hara and Slade (1998), and elsewhere. The rigorous results for two dimensions may be found in Kesten (1986b, c), who considers also the asymptotic properties of a random walk on the family tree of a critical branching process conditioned on non-extinction. See DeMasi, Ferrari, Goldstein, and Wick (1985, 1989) and Bezuidenhout and Grimmett (1997) for the central limit theorem associated with a random walk on the infinite open cluster of the supercritical percolation model.

Chayes, Chayes, and Durrett (1987) have studied the possible existence of an infinite open cluster in two dimensions when the probability that the edge e is open is $p_c + f(e)$, where $f(e) \approx |e|^{-\lambda}$ and $|e|$ is the distance from the origin to the midpoint of e. They show that the existence or not of an infinite open cluster depends on whether $\lambda < \nu^{-1}$ or $\lambda > \nu^{-1}$.

See Chayes and Chayes (1987a), Menshikov (1987b), and the references therein for further discussion of the backbone of the infinite open cluster.

Hara and Slade (1998) have studied the scaling limit of large finite clusters when $p = p_c$ and d is large. They have shown that the scaled two- and three-point connectivity functions correspond to those of so called 'integrated super-Brownian excursion'. See the notes for Section 10.3.

'Conformal invariance' has provided another approach to critical percolation when $d = 2$. This beautiful circle of ideas has provoked a detailed and quantitative picture which calls out for rigorous verification. See Cardy (1992), Langlands, Pouliot, and Saint-Aubin (1994), Aizenman (1995, 1997, 1998), Benjamini and Schramm (1998a), and Grimmett (1997, pp. 233–235). Roughly speaking, it is proposed that the large-scale structure of critical percolation clusters is invariant under conformal mappings of $\mathbb{R}^2$. This proposal gives rise to an exact formula for crossing probabilities known as 'Cardy's formula'; see the notes for Section 11.3. The work of Aizenman (1997, 1998) includes some detailed study of the incipient infinite cluster in two dimensions.

Chapter 10

Near the Critical Point: Rigorous Results

10.1 Percolation on a Tree

It is the presence of circuits in $\mathbb{L}^d$ which causes difficulties in exact calculations: there are many different paths joining two specified vertices, and pairs of such paths generally have edges in common. In attempting to understand critical phenomena, it is usual to tackle first the corresponding problem on a rather special lattice, being a lattice which is devoid of circuits; we do this in the hope that the ensuing calculation will be simple but will help in the development of insight into more general situations. Such a lattice is called a *tree*, and it is to percolation on trees that this section is devoted.

Let T be the labelled binary tree, illustrated in Figure 10.1. We label the vertices of T with sequences of 1's and 2's in the following way. We label the *root* of T by $\varnothing$, the empty sequence. We label the two children of the vertex labelled $\lambda_1\lambda_2\dots\lambda_n$ with the labels $\lambda_1\lambda_2\dots\lambda_n 1$ and $\lambda_1\lambda_2\dots\lambda_n 2$. Any vertex whose label is a sequence of length n is said to belong to the nth generation of T.

Let p satisfy $0 < p < 1$. We declare each edge of T to be *open* with probability p and *closed* otherwise, independently of all other edges. We call the resulting process *(bond) percolation on* T, and it is on this process that we shall concentrate.

We note that T is not a regular tree, since all vertices have degree 3 except the root which has degree 2. It is customary to add another branch from the root, isomorphic to the two which are already present, in order to equalize the degrees of the vertices. Such an addition has no effect on the critical exponents of the percolation process on T; since it is primarily in these exponents that we shall be interested, we shall not bother to do this. The critical exponents are similarly insensitive to the common degree of the vertices of T; thus we lose no substantial

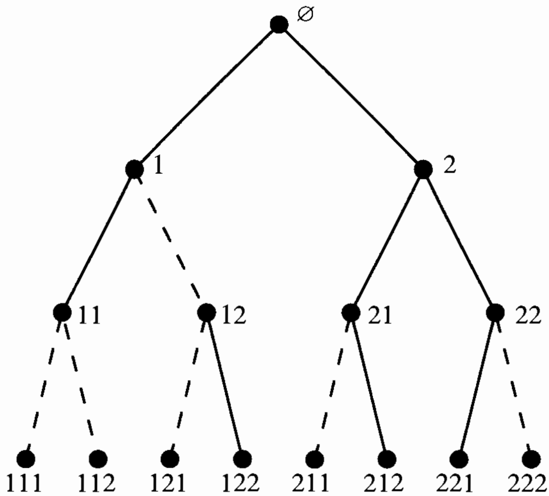


Figure 10.1. Four generations of the labelled binary tree, the unbroken edges being open and the broken edges being closed.

degree of generality by restricting our attention to the *binary* tree rather than considering the 'r-ary' tree in which each vertex has r children.

Such trees are sometimes called *Bethe lattices* after Hans Bethe. Mathematical physicists sometimes use the term 'mean field theory' to describe the statistical mechanics of interaction processes on trees; this is not the only use of the term.

There is a sense in which percolation on T corresponds to percolation on $\mathbb{L}^d$ in the limit as $d \to \infty$, and such an interpretation contributes to our understanding of percolation in high dimensions. We may think of T as being embedded in $\mathbb{L}^\infty$, with each edge joining an nth generation vertex to an $(n+1)$th generation vertex of T lying parallel to the nth coordinate axis of $\mathbb{L}^\infty$. In this embedding, the vertex of T with label $\lambda_1\lambda_2\ldots\lambda_n$ corresponds to the vertex $(x_1, x_2, \ldots, x_n, 0, 0, \ldots)$ where $x_i = -1$ if $\lambda_i = 1$ and $x_i = 1$ if $\lambda_i = 2$. In performing such an embedding we need to be careful when calculating distances between vertices. The natural distance function on T is defined as follows: the distance $\delta(x, y)$ is the number of edges in the unique path of T joining x to y. Thus $\delta(\varnothing, x)$ equals the generation number of x. With T embedded in $\mathbb{L}^\infty$, it is not unnatural to consider instead the Euclidean distance

$$(10.1) \qquad d(\varnothing, x) = \sqrt{\sum_{i=1}^{\infty} x_i^2},$$

where $(x_1, x_2, \ldots)$ is the vertex of $\mathbb{L}^\infty$ corresponding to x. It is not difficult to see that $d(\varnothing, x) = \sqrt{\delta(\varnothing, x)}$, and this difference between d and δ should be held in our minds.

Exact calculations are possible for percolation on a tree. We shall show how to perform these, but in advance of doing this we present a preview of our conclusions.

It is easy to see that the critical probability of percolation on T is given by $p_c = \frac{1}{2}$. We may calculate the critical exponents of the process exactly, and we shall see that

$$\alpha = -1,\ \beta = 1,\ \gamma = 1,\ \delta = 2,\ \Delta = 2,\ \rho = \tfrac{1}{2}.$$

The definitions in Section 9.1 of η and the correlation length $\xi(p)$ are not suitable for percolation on a tree. At the end of this section, we discuss alternative definitions of these quantities and calculate that $\eta = 0$ and $\nu = \frac{1}{2}$. The numerical values of ρ, η, and ν depend on which distance function is used in their definitions. The values above are calculated using the 'Euclidean' distance function $d(\varnothing, x)$; had we used $\delta(\varnothing, x)$ instead, they would each have been 1.

We note that the numerical values of these critical exponents satisfy the scaling relations

$$2 - \alpha = \gamma + 2\beta = \beta(\delta + 1),$$
$$\Delta = \beta\delta.$$

As we have remarked earlier, substitution of these numerical values into the hyperscaling relations $d\nu = 2 - \alpha$, $d\rho = \delta + 1$, yields $d = 6$, and this may be seen as evidence that the critical exponents for percolation on $\mathbb{L}^d$ take on the corresponding values for percolation on a tree when $d \geq 6$.

We write C for the open cluster of T containing the root $\varnothing$. Clearly, C is the family tree of a Galton–Watson branching process with a single progenitor and family sizes having the binomial distribution with parameters 2 and p. The cluster C is finite if and only if this branching process becomes extinct, and the probability of this is the smallest non-negative root of the equation $s = G(s)$, where $G(s) = (1 - p + ps)^2$ is the probability generating function of a typical family size. It is easy to check that this root equals 1 if $p \leq \frac{1}{2}$ and equals $\{(1-p)/p\}^2$ if $p \geq \frac{1}{2}$. We have proved that $\theta(p) = P_p(|C| = \infty)$ is given by

$$\theta(p) = \begin{cases} 0 & \text{if } p \leq \frac{1}{2}, \\ 1 - \left(\dfrac{1-p}{p}\right)^2 & \text{if } p \geq \frac{1}{2}. \end{cases} \tag{10.2}$$

Thus the critical probability p_c of percolation on T equals $\frac{1}{2}$, and the percolation probability $\theta(p)$ may be expressed in the closed form (10.2). We are interested here in critical exponents, and so we differentiate the function $1 - \{(1-p)/p\}^2$ at $p = \frac{1}{2}$ to obtain

$$\theta(p) \sim 8(p - \tfrac{1}{2}) \qquad \text{as } p \downarrow \tfrac{1}{2}. \tag{10.3}$$

We compare this with the hypothesis that $\theta(p) \approx (p - p_c)^\beta$ for some β as $p \downarrow p_c$ to arrive at the following proposition.

(10.4) Proposition. *For percolation on the binary tree, the critical exponent* β *equals* 1.

We shall see that all critical exponents may be calculated exactly and rigorously in the case of percolation on T. We consider γ next. Suppose that $p < p_c \ (= \frac{1}{2})$ and write $\chi(p) = E_p|C|$ as usual. The vertex x of T lies in C if and only if every edge is open in the unique path of the tree joining $\varnothing$ to x. There are 2^n vertices in the nth generation, so that

(10.5)
$$\begin{aligned}\chi(p) &= \sum_{n=0}^{\infty} 2^n p^n \\ &= (1-2p)^{-1} \qquad \text{if } p < \tfrac{1}{2} \\ &= \tfrac{1}{2}(\tfrac{1}{2} - p)^{-1},\end{aligned}$$

which we compare with the hypothesis that $\chi(p) \approx (p_c - p)^{-\gamma}$ as $p \uparrow p_c$, to deduce that $\gamma = 1$. The complete hypothesis involving γ asserts that $\chi^f(p) \approx |p - p_c|^{-\gamma}$ as $p \to p_c$, and we have verified this only in the case $p < p_c$. We therefore turn our attention to the case $p > p_c$, and recall that

$$\chi^f(p) = E_p(|C|;\ |C| < \infty),$$

the mean size of the family tree of the branching process on the event that the process becomes extinct. Let X_n be the number of nth generation descendants of $\varnothing$ in the branching process. The sequence $X = (X_n : n \geq 0)$ is a Markov chain with transition probabilities

(10.6)
$$P_p\big(X_{n+1} = j \mid X_n = i\big) = \binom{2i}{j} p^j (1-p)^{2i-j}.$$

Let $\overline{X} = (\overline{X}_n : n \geq 0)$ be the sequence of generation sizes of the branching process conditioned on the event E that the process becomes extinct. It is not difficult to check that $\overline{X}$ is a Markov chain whose transition probabilities are given by the following calculation (in which 'P_p' is used to denote the appropriate probability measure):

$$\begin{aligned}P_p\big(\overline{X}_{n+1} = j \mid \overline{X}_n = i\big) &= \frac{P_p(\overline{X}_{n+1} = j,\ \overline{X}_n = i)}{P_p(\overline{X}_n = i)} \\ &= \frac{P_p(X_{n+1} = j,\ X_n = i,\ E)/P_p(E)}{P_p(X_n = i,\ E)/P_p(E)} \\ &= \frac{P_p(X_{n+1} = j,\ E \mid X_n = i)P_p(X_n = i)}{P_p(E \mid X_n = i)P_p(X_n = i)} \\ &= \frac{P_p(X_{n+1} = j \mid X_n = i)P_p(E)^j}{P_p(E)^i} \qquad \text{for } i, j \geq 0.\end{aligned}$$

We substitute from (10.6) and use the fact that

$$P_p(E) = 1 - \theta(p) = \left(\frac{1-p}{p}\right)^2$$

to find that

$$\begin{aligned} P_p\big(\overline{X}_{n+1} = j \,\big|\, \overline{X}_n = i\big) &= \binom{2i}{j} p^j (1-p)^{2i-j} \left(\frac{1-p}{p}\right)^{2(j-i)} \\ &= \binom{2i}{j} (1-p)^j p^{2i-j}, \end{aligned}$$

which has the same form as (10.6) with p replaced by $1-p$. We have thus shown that the distribution of the branching process, conditional on extinction, is identical to that of a subcritical branching process having parameter $1-p$. In particular,

$$\begin{aligned} \chi^{\mathrm{f}}(p) &= E_p(|C|;\,|C|<\infty) \\ &= E_p\big(|C| \,\big|\, |C|<\infty\big) P_p(|C|<\infty) \\ &= \chi(1-p)\big(1-\theta(p)\big) \\ &= \tfrac{1}{2}(p-\tfrac{1}{2})^{-1}\left(\frac{1-p}{p}\right)^2 \end{aligned} \tag{10.7}$$

by (10.2) and (10.5). Thus $\chi^{\mathrm{f}}(p) \approx (p-p_{\mathrm{c}})^{-1}$ when $p > p_{\mathrm{c}}$, and we have proved the next proposition.

(10.8) Proposition. *For percolation on the binary tree, the critical exponent* γ *equals* 1.

Whilst on the subject of the mean size of C, we recall the gap exponent Δ given by

$$\frac{E_p(|C|^{k+1};\,|C|<\infty)}{E_p(|C|^k;\,|C|<\infty)} \approx |p-p_{\mathrm{c}}|^{-\Delta} \qquad \text{as } p \to p_{\mathrm{c}} \tag{10.9}$$

for $k \geq 1$. We may suppose that $p < p_{\mathrm{c}}\ (=\frac{1}{2})$, since if $p > \frac{1}{2}$ then the discussion prior to (10.7) gives that

$$E_p(|C|^k;\,|C|<\infty) = E_{1-p}(|C|^k)P_p(|C|<\infty). \tag{10.10}$$

We suppose therefore that $p < \frac{1}{2}$ and $k \geq 1$. Then

$$E_p(|C|^k;\,|C|<\infty) = E_p(|C|^k) = \sum_{n=1}^{\infty} n^k P_p(|C|=n). \tag{10.11}$$

There is more than one way to discover how this behaves as p approaches $\frac{1}{2}$ from beneath, and here is such a way. It is not difficult to calculate $P_p(|C| = n)$ exactly: either use probability generating functions and the natural recursion, or read Durrett (1985a, p. 429) and the references therein, to find that

$$P_p(|C| = n) = \frac{1}{n}\binom{2n}{n-1}p^{n-1}(1-p)^{n+1}. \tag{10.12}$$

We write $\varepsilon = \frac{1}{2} - p$ and substitute (10.12) into (10.11) to find that

$$E_p(|C|^k) = \sum_{n=1}^{\infty} n^{k-1}\binom{2n}{n-1}(\tfrac{1}{2}-\varepsilon)^{n-1}(\tfrac{1}{2}+\varepsilon)^{n+1}. \tag{10.13}$$

We have from Stirling's formula that

$$\binom{2n}{n-1} = \frac{n}{n+1}\binom{2n}{n} \sim \frac{4^n}{\sqrt{\pi n}} \qquad \text{as } n \to \infty.$$

We approximate the right side of (10.13) by an integral and we make the substitution $x = -n\log(1-4\varepsilon^2)$ to obtain

(10.14)

$$\begin{aligned} A\int_1^\infty n^{k-1}\frac{(1-4\varepsilon^2)^n}{\sqrt{n}}\,dn &\sim A\big\{-\log(1-4\varepsilon^2)\big\}^{-k+\frac{1}{2}}\int_0^\infty x^{k-\frac{3}{2}}e^{-x}\,dx \\ &\sim A(4\varepsilon^2)^{-k+\frac{1}{2}}\Gamma(k-\tfrac{1}{2}), \end{aligned}$$

as $\varepsilon \downarrow 0$, where A is a constant and Γ is the gamma function. This asymptotic analysis is rather approximate, but it is not difficult to make it rigorous when $k \geq 1$. We compare the last formula with (10.9), and obtain the next result.

(10.15) Proposition. *For percolation on the binary tree, the critical exponent* Δ *equals* 2.

The nature of the above calculation suggests looking for more powerful asymptotic results as $p \to \frac{1}{2}$ than merely those concerning the moments of $|C|$. One way to do this is to define a random variable Y_p having mass function

$$P(Y_p = n) = \frac{nP_p(|C| = n)}{\sum_m mP_p(|C| = m)} = \frac{nP_p(|C| = n)}{\chi^{\mathrm{f}}(p)}. \tag{10.16}$$

Some people think of Y_p as having approximately the distribution of a 'typical' finite open cluster, but the intuition behind this is not entirely convincing, and we omit it.

It follows from the discussion leading to (10.7) that Y_p and Y_{1-p} have the same distribution. It is a reasonably elementary exercise in branching processes and asymptotic analysis to show that $|p-\frac{1}{2}|^2 Y_p$ converges in distribution as $p \to \frac{1}{2}$, with the limit distribution being gamma with parameters $\frac{1}{2}$ and 4. That is to say,

$$(10.17)\qquad P\big(|p-\tfrac{1}{2}|^2 Y_p \le y\big) \to \int_0^y \frac{2}{\sqrt{\pi}} x^{-1/2} e^{-4x}\,dx$$

as $p \to \frac{1}{2}$. A similar calculation shows that, for $k \ge 1$,

$$(10.18)\qquad \begin{aligned} E_p\big(|C|^k; |C|<\infty\big) &= E(Y_p^{k-1})\chi^{\mathrm{f}}(p) && \text{by (10.16)}\\ &\sim A_k|p-\tfrac{1}{2}|^{-2(k-1)}\chi^{\mathrm{f}}(p) && \text{as in (10.17)}\\ &\sim \tfrac{1}{2}A_k|p-\tfrac{1}{2}|^{1-2k} && \text{by (10.7)}\end{aligned}$$

for some constant A_k depending on k, in agreement with (10.14).

The critical exponent α is defined by $\kappa'''(p) \approx |p-p_c|^{-1-\alpha}$ as $p \to p_c$, where $\kappa(p) = E_p(|C|^{-1})$. We substitute $k=-1$ into (10.13) to find that

$$(10.19)\qquad \kappa(p) = \frac{1+2\varepsilon}{1-2\varepsilon}\sum_{n=1}^{\infty}\frac{1}{n^2}\binom{2n}{n-1}\frac{1}{4^n}(1-4\varepsilon^2)^n,$$

where $\varepsilon = \frac{1}{2}-p$. Differentiating with respect to ε, we discover that the infinite series in (10.19) is infinitely differentiable term by term, except possibly at $\varepsilon = 0$. It is now a minor calculation somewhat similar to (10.14) to show that κ''' is bounded on a set of the form $(\frac{1}{2}-\zeta, \frac{1}{2}) \cup (\frac{1}{2}, \frac{1}{2}+\zeta)$ for some $\zeta > 0$, and thus $\alpha = -1$. It turns out that $\kappa'''(p)$ has a jump discontinuity at $p = \frac{1}{2}$.

(10.20) Proposition. *For percolation on the binary tree, the critical exponent* α *equals* -1.

Little extra work is needed to calculate the critical exponent δ. We have from (10.12) and Stirling's formula that

$$P_{1/2}(|C| = n) \sim \frac{1}{\sqrt{\pi n^3}},$$

giving that $\delta = 2$. For the critical exponent ρ, we argue as follows. If $p = \frac{1}{2}$, the corresponding branching process is critical, and a version of the classical result of Kolmogorov (see Harris (1963, p. 21)) implies that

$$(10.21)\qquad P_{1/2}\big(\mathrm{rad}(C) \ge n\big) \sim \frac{4}{n} \qquad \text{as } n \to \infty,$$

where $\mathrm{rad}(C) = \max\{\delta(\varnothing, x) : x \in C\}$; it is an interesting exercise to derive this from first principles. Mathematical physicists prefer to use the 'Euclidean' distance $d(\varnothing, x) = \sqrt{\delta(\varnothing, x)}$. We rewrite (10.21) as

$$P_{1/2}\left(\sqrt{\mathrm{rad}(C)} \ge n\right) \sim \frac{4}{n^2} \qquad \text{as } n \to \infty,$$

and compare this with the hypothesis that $P_{p_c}\big(\sqrt{\mathrm{rad}(C)} = n\big) \approx n^{-1-1/\rho}$, to obtain $\rho = \frac{1}{2}$.

(10.22) Proposition. *For bond percolation on the binary tree, the critical exponents δ and ρ satisfy $\delta = 2$, $\rho = \frac{1}{2}$.*

There remain the exponents η and ν, defined in Section 9.1 by

$$P_{p_c}(0 \leftrightarrow x) \approx |x|^{2-d-\eta}, \quad \xi(p) \approx |p - p_c|^{-\nu}. \tag{10.23}$$

This definition of η is not suitable for percolation on a tree, since

$$P_{1/2}(\varnothing \leftrightarrow x) = \left(\tfrac{1}{2}\right)^{\delta(\varnothing,x)},$$

which does not decay as a power of $\delta(\varnothing, x)$. A similar problem arises for ν, since the definition of the correlation length $\xi(p)$ is not suitable for percolation on trees. We complete this section with a discussion of alternative definitions of η and $\xi(p)$ suitable for the binary tree, and we begin with η. Rather than basing our approach upon the connectivity function $\tau_{p_c}(0, x)$, we work with the number of vertices within distance n of the origin which are joined to the origin by open paths. Let $S(n)$ be the ball with radius n and centre at the origin. A rough calculation for percolation on $\mathbb{L}^d$ suggests

$$\begin{aligned} E_{p_c}|C \cap S(n)| &= \sum_{x:|x|\le n} \tau_{p_c}(0, x) \\ &\simeq \sum_{m=0}^{n} m^{d-1} m^{2-d-\eta} \qquad \text{by (10.23)} \\ &\simeq n^{2-\eta}, \end{aligned} \tag{10.24}$$

and we take this formula as our template for the corresponding calculation for the binary tree. In the latter case, we have that

$$\begin{aligned} E_{p_c}|C \cap S(n)| &= \sum_{x:\delta(\varnothing,x)\le n} P_{1/2}(\varnothing \leftrightarrow x) = \sum_{m=0}^{n} 2^m \left(\tfrac{1}{2}\right)^m \\ &\sim n, \end{aligned}$$

implying that

$$E_{p_c}\left|\{x : \varnothing \leftrightarrow x \text{ and } d(\varnothing, x) \le n\}\right| = E_{p_c}|C \cap S(n^2)| \sim n^2,$$

and thence $\eta = 0$ by comparison with (10.24).

As remarked in Section 9.1, another possible definition of correlation length is the quantity $\varphi(p)$ given by

$$\varphi(p) = \sqrt{\frac{1}{\chi^{\mathrm{f}}(p)} \sum_x |x|^2 \tau_p^{\mathrm{f}}(0, x)}.$$

Once again there is ambiguity over the choice of the distance function $|x|$. Following the preference of mathematical physicists, we substitute $\sqrt{\delta(\varnothing, x)}$ for $|x|$ and obtain

$$\varphi(p) = \sqrt{\frac{1}{\chi^{\mathrm{f}}(p)} \sum_{n=0}^{\infty} n 2^n \zeta_p(n)},$$

where $\zeta_p(n)$ is the probability that there exists an open path in T joining the root $\varnothing$ to a given vertex in the nth generation, but that the branching process with progenitor $\varnothing$ is extinct. It is not difficult to show that

$$\zeta_p(n) = \big(1 - \theta(p)\big)\zeta_{1-p}(n) \quad \text{if } p > \tfrac{1}{2}$$

as in the discussion prior to (10.7), and furthermore

$$\chi^{\mathrm{f}}(p) = \big(1 - \theta(p)\big)\chi(1 - p) \quad \text{if } p > \tfrac{1}{2}.$$

Thus $\varphi(p) = \varphi(1 - p)$, so that the behaviour of $\varphi(p)$ when p is near to $\frac{1}{2}$ is determined by its behaviour for $p < \frac{1}{2}$. Suppose then that $p < \frac{1}{2}$. We have that $\zeta_n(p) = p^n$, giving that

$$\begin{aligned}
(10.25) \qquad \varphi(p) &= \sqrt{\frac{1}{\chi^{\mathrm{f}}(p)} \sum_{n=0}^{\infty} n(2p)^n} \\
&= \sqrt{\frac{p}{\frac{1}{2} - p}} && \text{by (10.5)} \\
&\approx (\tfrac{1}{2} - p)^{-1/2} && \text{as } p \uparrow \tfrac{1}{2}.
\end{aligned}$$

(10.26) Proposition. *For percolation on the binary tree, the critical exponents η and ν satisfy $\eta = 0$, $\nu = \frac{1}{2}$.*

10.2 Inequalities for Critical Exponents

We return to the muddier waters of percolation on $\mathbb{L}^d$ when $d \geq 2$. As we have recounted earlier, it is believed that the critical exponents take on the corresponding values for percolation on a tree when d is sufficiently large, and there is reason to suppose that this occurs first when $d = 6$. We sketch an approach to proving this in Section 10.3. At present, only certain results for $d \geq 19$ are known rigorously. Whereas no exact formulae are known for critical exponents on $\mathbb{L}^d$ for general d (indeed, the very existence of critical exponents remains unproven in general), we have at our disposal certain rigorous inequalities which are valid for all $d \geq 2$.

Consider for example the exponent β for the percolation probability $\theta(p)$. We saw in Theorems (5.8) and (5.48) that

$$\theta(p) - \theta(p_c) \geq a(p - p_c)$$

for some a (> 0) and for all small positive values of $p - p_c$. It follows that, if $\theta(p) \approx (p - p_c)^\beta$ for $p > p_c$, then $\beta \leq 1$.

(10.27) Proposition. *If the critical exponent β exists, then $\beta \leq \beta_T$, where $\beta_T = 1$ is the corresponding exponent for percolation on the binary tree.*

Thus β is bounded on one side by its value for a tree; it is conjectured that equality is valid here if $d \geq 6$. Similar inequalities are expected for the other critical exponents, and we describe some such results here. We assume throughout this section that $\theta(p_c) = 0$, although this is not essential for all our conclusions.

(10.28) Proposition. *If the critical exponent γ exists, then $\gamma \geq \gamma_T$, where $\gamma_T = 1$ is the corresponding exponent for percolation on the binary tree.*

(10.29) Proposition. *If the critical exponent δ exists, then $\delta \geq \delta_T$, where $\delta_T = 2$ is the corresponding exponent for percolation on the binary tree. If, further, the critical exponent β exists, then $\beta \geq 2/\delta$.*

We note that $\beta_T = 2/\delta_T$; many mathematical physicists express this by saying that 'the inequality $\beta \geq 2/\delta$ saturates in mean field'. This inequality is a strengthening of the fact that $\delta \geq 2$, since $\beta \leq 1$.

Many other inequalities for critical exponents are known, and we shall not list them here. For references, see the notes at the end of this chapter.

Proof of Proposition (10.28). We follow Aizenman and Newman (1984). First we express $\chi(p)$ as the sum of the two-point connectivities:

(10.30) $$\chi(p) = \sum_{x \in \mathbb{Z}^d} \tau_p(0, x) = \sum_{x \in \mathbb{Z}^d} P_p(0 \leftrightarrow x).$$

We should like to differentiate with respect to p and use Russo's formula, but there are two obstacles in that the summations are infinite and that the events $\{0 \leftrightarrow x\}$ depend on the states of infinitely many edges. In order to overcome this difficulty, we restrict ourselves for the moment to percolation on the finite box $B(n)$, and later we shall take the limit as $n \to \infty$. We write $\chi_n(p, v)$ for the mean number of vertices in the box $B(n)$ which are joined to the vertex v by open paths of $B(n)$, and we define

(10.31) $$\widehat{\chi}_n(p) = \max\{\chi_n(p, v) : v \in B(n)\}.$$

Now, $\chi_n(p, v) \leq \chi(p)$ for all $v \in B(n)$, giving that $\widehat{\chi}_n(p) \leq \chi(p)$. On the other hand,

(10.32) $$\widehat{\chi}_n(p) \geq \chi_n(p, 0) = \sum_{x \in B(n)} P_p(0 \leftrightarrow x \text{ in } B(n)).$$

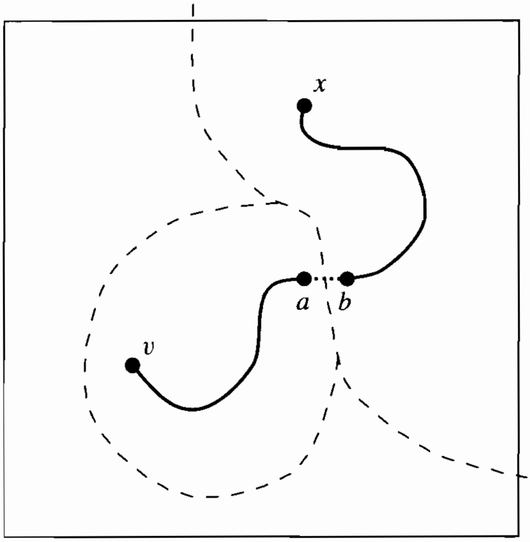


Figure 10.2. A sketch of the event that $e = \langle a, b\rangle$ is pivotal for $A_n(v, x)$. If e is open then $v \leftrightarrow x$ in $B(n)$, but not otherwise.

However,

$$P_p\big(0 \leftrightarrow x \text{ in } B(n)\big) \to P_p(0 \leftrightarrow x) \qquad \text{as } n \to \infty,$$

so that

$$\begin{aligned}\chi(p) &\geq \widehat{\chi}_n(p)\\ &\geq \sum_{x \in B(n)} P_p\big(0 \leftrightarrow x \text{ in } B(n)\big)\\ &\to \sum_{x \in \mathbb{Z}^d} P_p(0 \leftrightarrow x) \qquad \text{as } n \to \infty\\ &= \chi(p),\end{aligned}$$

by the bounded convergence theorem. Therefore,

$$\widehat{\chi}_n(p) \to \chi(p) \qquad \text{as } n \to \infty, \tag{10.33}$$

whether or not $\chi(p)$ is finite.

Let v be a vertex of $B(n)$. We shall concentrate upon $\chi_n(p, v)$, which we write in the form

$$\chi_n(p, v) = \sum_{x \in B(n)} P_p\big(v \leftrightarrow x \text{ in } B(n)\big).$$

By Russo's formula,

$$\frac{d}{dp}\chi_n(p, v) = \sum_{x\in B(n)} \sum_{e\in B(n)} P_p\big(e \text{ is pivotal for } A_n(v, x)\big), \tag{10.34}$$

where $A_n(v, x)$ is the event that there exists an open path from v to x in $B(n)$, and the second summation is over all edges e in $B(n)$. Writing $e = \langle a, b\rangle$, we have that, if e is pivotal for $A_n(v, x)$, there exists an open path in $B(n)$ from v to either a or b, and there exists a disjoint open path in $B(n)$ from x to the other endvertex of e. See Figure 10.2 for a sketch of the situation. Therefore,

(10.35)
$$\begin{aligned} &P_p\big(e \text{ is pivotal for } A_n(v, x)\big) \\ &\qquad \le P_p\big(A_n(v, a) \circ A_n(x, b)\big) + P_p\big(A_n(v, b) \circ A_n(x, a)\big) \\ &\qquad \le P_p\big(A_n(v, a)\big)P_p\big(A_n(x, b)\big) + P_p\big(A_n(v, b)\big)P_p\big(A_n(x, a)\big) \end{aligned}$$

by the BK inequality. We substitute this into (10.34) and sum over all vertices x and edges $e = \langle a, b\rangle$ to obtain

(10.36)
$$\begin{aligned} \frac{d}{dp}\chi_n(p, v) &\le \sum_{e=\langle a,b\rangle} \Big\{P_p\big(A_n(v, a)\big)\chi_n(p, b) + P_p\big(A_n(v, b)\big)\chi_n(p, a)\Big\} \\ &\le \widehat{\chi}_n(p) \sum_{e=\langle a,b\rangle} \Big\{P_p\big(A_n(v, a)\big) + P_p\big(A_n(v, b)\big)\Big\} \\ &\le 2d\widehat{\chi}_n(p)\chi_n(p, v) \\ &\le 2d\widehat{\chi}_n(p)^2. \end{aligned}$$

It would be convenient to take the limit as $n \to \infty$ in order to deduce that

$$\frac{d}{dp}\chi(p) \le 2d\chi(p)^2, \tag{10.37}$$

but certain difficulties arise in justifying the step $\chi_n'(p, v) \to \chi'(p)$. These difficulties may be overcome directly, but we shall argue differently here. We note that $\widehat{\chi}_n(p)$ is the maximum of a finite number of polynomial functions $\chi_n(p, v)$. Therefore $\widehat{\chi}_n(p)$ is differentiable except possibly for finitely many values of p. Also, whenever $\widehat{\chi}_n$ is differentiable, we have that

$$\begin{aligned} \frac{d}{dp}\widehat{\chi}_n(p) &\le \max_{v\in B(n)} \left\{\frac{d}{dp}\chi_n(p, v)\right\} \\ &\le 2d\widehat{\chi}_n(p)^2 \qquad \text{by (10.36).} \end{aligned} \tag{10.38}$$

To recapitulate, we have that $\widehat{\chi}_n(p)$ is continuous on $[0, 1]$ and satisfies

$$\widehat{\chi}_n(p)^{-2}\frac{d}{dp}\widehat{\chi}_n(p) \le 2d \tag{10.39}$$

except possibly for a finite set of values of p. We integrate (10.39) over the interval $[p, p_c]$ to obtain

$$\left[-\frac{1}{\widehat{\chi}_n(\pi)}\right]_p^{p_c} \leq 2d(p_c - p),$$

or

$$\frac{1}{\widehat{\chi}_n(p)} - \frac{1}{\widehat{\chi}_n(p_c)} \leq 2d(p_c - p). \tag{10.40}$$

We may at last allow the limit $n \to \infty$. As $n \to \infty$, it is the case that $\widehat{\chi}_n(p) \to \chi(p)$ by (10.33), and $\widehat{\chi}_n(p_c) \to \chi(p_c) = \infty$ by (6.52). Inequality (10.40) becomes

$$\chi(p) \geq \frac{1}{2d(p_c - p)} \qquad \text{for } p < p_c. \tag{10.41}$$

Comparison with the hypothesis $\chi(p) \approx (p_c - p)^{-\gamma}$ for $p < p_c$ yields $\gamma \geq 1$ as required. Note that we have used the fact that $\chi(p_c) = \infty$, proved in (6.52). This reference to earlier work is easily avoided by noting that (10.39) is valid for all $p \in [0, 1]$ at which $\widehat{\chi}_n$ is differentiable. We integrate (10.39) over the interval $[p, p_c + \varepsilon]$ where ε is small, to obtain

$$\frac{1}{\widehat{\chi}_n(p)} - \frac{1}{\widehat{\chi}_n(p_c + \varepsilon)} \leq 2d(p_c + \varepsilon - p).$$

Now $\widehat{\chi}_n(p_c + \varepsilon) \to \chi(p_c + \varepsilon) = \infty$ as $n \to \infty$, and ε was arbitrary, so that (10.41) is valid as before. □

Proof of Proposition (10.29). We give two proofs that $\delta \geq 2$. For the first, we follow Aizenman and Barsky (1987). The ground was prepared for this in Chapter 5, whence we extract Proposition (5.49): if p is such that $\chi^f(p) = \infty$, then

$$\begin{aligned} \theta(p, \psi) &= 1 - \sum_{n=1}^{\infty} (1 - \psi)^n P_p(|C| = n) \\ &\geq a\psi^{1/2} \end{aligned}$$

for some $a = a(p) > 0$ and all small positive values of ψ. We apply this with $p = p_c$, noting that $\chi^f(p_c) = \chi(p_c) = \infty$ if $\theta(p_c) = 0$, to obtain

$$\theta(p_c, \psi) \geq a\psi^{1/2} \tag{10.42}$$

for small positive ψ.

The inequality $\delta \geq 2$ follows by a standard application of the following Tauberian theorem for power series (see Feller (1971, p. 447)). Let $(q_m : m \geq 1)$ be a monotonic sequence of positive real numbers with the property that the series

$$Q(\psi) = \sum_{m=1}^{\infty} (1 - \psi)^m q_m$$

converges for $0 < \psi \le 1$. If L is a slowly varying function and $0 < \zeta < 1$ then the two relations

$$Q(\psi) \sim \psi^{-\zeta} L(\psi^{-1}) \qquad \text{as } \psi \downarrow 0$$

and

$$q_m \sim \frac{1}{\Gamma(\zeta)} m^{\zeta-1} L(m) \qquad \text{as } m \to \infty$$

are equivalent.

We may apply this theorem as follows. First, we note that

(10.43)

$$\begin{aligned}\sum_{m=1}^{\infty}(1-\psi)^m P_{p_c}(|C| \ge m) &= \sum_{m=1}^{\infty}\sum_{n=m}^{\infty}(1-\psi)^m P_{p_c}(|C|=n)\\ &= \frac{1-\psi}{\psi}\sum_{n=1}^{\infty}\{1-(1-\psi)^n\}P_{p_c}(|C|=n)\\ &= \frac{1-\psi}{\psi}\theta(p_c,\psi),\end{aligned}$$

where we have used the assumption that $\theta(p_c) = 0$ in the form

$$\sum_n P_{p_c}(|C| = n) = 1.$$

Let us assume that

(10.44)
$$P_{p_c}(|C| = m) \approx m^{-1-1/\delta} \qquad \text{as } m \to \infty$$

where $1 \le \delta < \infty$. For $0 < \varepsilon < \delta^{-1}$, there exists a constant A_1 such that

$$P_{p_c}(|C| \ge m) \le A_1 m^{\varepsilon - 1/\delta} \qquad \text{for } m \ge 1.$$

Applying the Tauberian theorem with $q_m = A_1 m^{\varepsilon-1/\delta}$, we find that

$$\sum_{m=1}^{\infty}(1-\psi)^m P_{p_c}(|C| \ge m) \le A_2 \psi^{-(1+\varepsilon-1/\delta)}$$

for some constant A_2 and all small positive ψ. By (10.43),

(10.45)
$$\theta(p_c, \psi) \le \frac{A_2}{1-\psi}\psi^{-\varepsilon+1/\delta}$$

for all small positive ψ, yielding in the light of (10.42) that $\delta^{-1} \le \frac{1}{2} + \varepsilon$. We now take the limit as $\varepsilon \downarrow 0$ to obtain that $\delta \ge 2$.

It is easy to see by the foregoing arguments that (10.44), taken together with the inequality $2 \le \delta < \infty$, implies that

$$\theta(p_c, \psi) \approx \psi^{1/\delta} \qquad \text{as } \psi \downarrow 0, \tag{10.46}$$

an expression which has been used by some as the formal definition of the critical exponent δ.

We follow Newman (1987c) for the rest of this proof. Let us write $\Pi_{nmb}(p)$ for the probability

$$\Pi_{nmb}(p) = a_{nmb} p^m (1-p)^b \tag{10.47}$$

that the origin belongs to an open cluster having n vertices, m edges, and b boundary edges; as usual, a_{nmb} is the number of such animals. It is easy to see that

$$\Pi_{nmb}(p-\varepsilon) = \left(1 - \frac{\varepsilon}{p}\right)^m \left(1 + \frac{\varepsilon}{1-p}\right)^b \Pi_{nmb}(p) \tag{10.48}$$

for any small ε, positive or negative. We wish to relate β to δ, which is to say that we wish to relate $\theta(p_c + \varepsilon)$ to $\theta(p_c, \psi)$, for $\varepsilon, \psi > 0$. We argue as follows. Suppose that $\varepsilon > 0$. Then

$$P_{p_c - \varepsilon}(|C| < \infty) = \sum_{n,m,b} \Pi_{nmb}(p_c - \varepsilon) = 1,$$

so that

$$\begin{aligned}
\theta(p_c + \varepsilon) &= 1 - \sum_{n,m,b} \Pi_{nmb}(p_c + \varepsilon) \\
&= 1 - \left(\sum_{n,m,b} \Pi_{nmb}(p_c - \varepsilon)\right)\left(\sum_{n,m,b} \Pi_{nmb}(p_c + \varepsilon)\right) \\
&= 1 - E_{p_c}\left\{\left(1 - \frac{\varepsilon}{p_c}\right)^{|C_e|}\left(1 + \frac{\varepsilon}{1-p_c}\right)^{|\Delta C|}; |C| < \infty\right\} \\
&\qquad \times E_{p_c}\left\{\left(1 + \frac{\varepsilon}{p_c}\right)^{|C_e|}\left(1 - \frac{\varepsilon}{1-p_c}\right)^{|\Delta C|}; |C| < \infty\right\}
\end{aligned}$$

by (10.48), where C_e is the set of edges of C and ΔC is its external boundary. We use the Cauchy–Schwarz inequality to deduce that

$$\theta(p_c + \varepsilon) \le$$
$$1 - \left[E_{p_c}\left\{\left(1 - \frac{\varepsilon^2}{p_c^2}\right)^{|C_e|/2}\left(1 - \frac{\varepsilon^2}{(1-p_c)^2}\right)^{|\Delta C|/2}; |C| < \infty\right\}\right]^2.$$

However, $|C_e| \le 2d|C|$ and $|\Delta C| \le 2d|C|$ from (4.14) and (4.15), and it follows from the definition of $\theta(p, \psi)$ that

$$\theta(p_c + \varepsilon) \le 1 - \{1 - \theta(p_c, \psi(\varepsilon))\}^2, \tag{10.49}$$

where

$$\begin{aligned} \psi(\varepsilon) &= 1 - \left\{\left(1 - \frac{\varepsilon^2}{p_c^2}\right)\left(1 - \frac{\varepsilon^2}{(1-p_c)^2}\right)\right\}^d \\ &= A_1\varepsilon^2 + o(\varepsilon^2) \qquad \text{as } \varepsilon \downarrow 0 \end{aligned} \tag{10.50}$$

for some positive constant A_1. Now $\psi(\varepsilon) \to 0$ as $\varepsilon \downarrow 0$, and it follows that $\theta(p_c, \psi(\varepsilon)) \to 0$ also, giving from (10.49) that

$$\theta(p_c + \varepsilon) \le 2\big(1 + o(1)\big)\theta(p_c, \psi(\varepsilon)) \qquad \text{as } \varepsilon \downarrow 0. \tag{10.51}$$

We have from Theorem (5.8) that $\theta(p_c + \varepsilon) \ge a\varepsilon$ for some a (> 0) and all small ε, and thus (10.50) and (10.51) imply that

$$\theta(p_c, \psi) \ge A_2 \psi^{1/2}$$

for some positive constant A_2 and all small positive values of ψ. We compare this with (10.46) and deduce that $\delta \ge 2$, if δ exists.

The inequality $\beta \ge 2/\delta$ follows similarly: if $\theta(p_c, \psi) \approx \psi^{1/\delta}$ and $\theta(p_c + \varepsilon) \approx \varepsilon^\beta$ then (10.50) and (10.51) yield the required inequality. □

10.3 Mean Field Theory

The expression 'mean field' is often used by mathematical physicists, and it permits several interpretations depending on context. A narrow interpretation of the term 'mean field theory' for percolation involves trees rather than lattices. As we saw in Section 10.1, for percolation on a regular tree, it is quite easy to perform exact calculations of many quantities including the numerical values of critical exponents.

Turning to percolation on $\mathbb{L}^d$, it is known that several critical exponents agree with those of a regular tree when d is sufficiently large. This is believed to hold if and only if $d \ge 6$, but progress so far assumes that $d \ge 19$. In the following theorem, taken from Hara and Slade (1994), we write $f(x) \asymp g(x)$ if there exist positive constants c_1, c_2 such that $c_1 f(x) \le g(x) \le c_2 f(x)$ for all x close to a limiting value.

(10.52) Theorem. Mean field behaviour. *If* $d \geq 19$,

(10.53) $$\theta(p) \asymp (p - p_c)^1 \quad \text{as } p \downarrow p_c,$$

(10.54) $$\chi(p) \asymp (p_c - p)^{-1} \quad \text{as } p \uparrow p_c,$$

(10.55) $$\xi(p) \asymp (p_c - p)^{-1/2} \quad \text{as } p \uparrow p_c,$$

(10.56) $$\frac{\chi^f_{k+1}(p)}{\chi^f_k(p)} \asymp (p_c - p)^{-2} \quad \text{as } p \uparrow p_c, \text{ for } k \geq 1.$$

Note the strong form of the asymptotic relation $\asymp$, and the identification of the critical exponents β, γ, ν, Δ. The proof of Theorem (10.52) centres on a property known as the 'triangle condition'. Let $d \geq 2$. We define the 'triangle function'

(10.57) $$T(p) = \sum_{x,y \in \mathbb{Z}^d} P_p(0 \leftrightarrow x) P_p(x \leftrightarrow y) P_p(y \leftrightarrow 0),$$

and we introduce the following condition:

(10.58) *Triangle condition*: $\qquad T(p_c) < \infty.$

The triangle condition was first introduced by Aizenman and Newman (1984), who showed that it implies (10.54). Subsequently other authors showed that the triangle condition implies similar asymptotics for other quantities. It was Hara and Slade (1990) who verified the triangle condition for large d, exploiting a technique known as the 'lace expansion'.

The triangle function (10.57) involves convolutions, and it is therefore natural to use Fourier transforms, defined as follows. If $f : \mathbb{Z}^d \to \mathbb{R}$ is absolutely summable, we define

$$\widehat{f}(\theta) = \sum_{x \in Z^d} f(x) e^{i\theta \cdot x} \quad \text{for } \theta = (\theta_1, \theta_2, \ldots, \theta_d) \in [-\pi, \pi]^d,$$

where $\theta \cdot x = \sum_{j=1}^d \theta_j x_j$. If f is symmetric (that is, if $f(x) = f(-x)$ for all $x \in \mathbb{Z}^d$), its Fourier transform $\widehat{f}$ takes values in the reals; in particular, the Fourier transform $\widehat{\tau}_p$ of the two-point connectivity function $\tau_p(0, x) = P_p(0 \leftrightarrow x)$ is real.

The function τ_p is positive definite, in that

$$\begin{aligned} \sum_{v,w \in Z^d} f(v) \tau_p(v, w) \overline{f(w)} &= E_p\Bigl(\sum_{v,w} f(v) I_{\{v \leftrightarrow w\}} \overline{f(w)}\Bigr) \\ &= E_p\Bigl(\sum_C \Bigl|\sum_{x \in C} f(x)\Bigr|^2\Bigr) \\ &\geq 0 \end{aligned}$$

for any absolutely summable function $f : \mathbb{Z}^d \to \mathbb{C}$; the penultimate summation is over all open clusters C. It follows by Bochner's theorem (see Reed and Simon (1975, p. 13) or Rudin (1962, p. 19)) that

$$\widehat{\tau}_p(\theta) \geq 0 \quad \text{for } \theta \in [-\pi, \pi]^d. \tag{10.59}$$

Let

$$Q(u) = \sum_{v,w \in \mathbb{Z}^d} \tau_p(0, v)\tau_p(v, w)\tau_p(w, u), \quad u \in \mathbb{Z}^d.$$

We have that $\widehat{Q}(\theta) = \widehat{\tau}_p(\theta)^3$, whence, by the Fourier inversion theorem,

$$Q(u) = (2\pi)^{-d} \int_{[-\pi,\pi]^d} \widehat{\tau}_p(\theta)^3 e^{-i\theta \cdot u}\, d\theta, \tag{10.60}$$

and, in particular,

$$T(p) = Q(0) = (2\pi)^{-d} \int_{[-\pi,\pi]^d} \widehat{\tau}_p(\theta)^3\, d\theta.$$

We present no full proof of Theorem (10.52) here, pleading two reasons. First, such a proof would be long and complicated. Secondly, we are unable to do better than is already contained in the existing literature. Instead, we (nearly) prove (10.54) assuming the triangle condition, and then we present a very brief discussion of the Hara–Slade verification of the triangle condition for large d.

(10.61) Theorem. *If $d \geq 2$ and $T(p_c) < \infty$, then*

$$\chi(p) \asymp (p_c - p)^{-1} \quad \text{as } p \uparrow p_c.$$

Proof. This is taken from Aizenman and Newman (1984); see also Hara and Slade (1994). The required lower bound for $\chi(p)$ was given in (10.41), and we show here how to obtain a corresponding upper bound. For the sake of brevity, we omit the details of the finite-box approximation which is necessary for this, instead referring the reader to the proof of (10.41).

By (ab)use of Russo's formula applied to the sum

$$\chi(p) = \sum_{x \in \mathbb{Z}^d} \tau_p(0, x),$$

we have as in (10.34) that

$$\frac{d\chi}{dp} = \frac{d}{dp} \sum_{x \in \mathbb{Z}^d} \tau_p(0, x) = \sum_{x \in \mathbb{Z}^d} \sum_{e \in \mathbb{E}^d} P_p\big(e \text{ is pivotal for } \{0 \leftrightarrow x\}\big). \tag{10.62}$$

Unlike in Section 10.2, we shall use this formula to derive the required *upper* bound for $\chi(p)$.

Let $e = \langle a, b\rangle$ in (10.62), and change variables $(x \mapsto x - a)$ in the summation to obtain that

$$\frac{d\chi}{dp} = \sum_{x,y} \sum_{|u|=1} P_p\big(0 \leftrightarrow x,\ u \leftrightarrow y \text{ off } C_u(x)\big) \tag{10.63}$$

where the second summation is over all (positive and negative) unit vectors u of $\mathbb{Z}^d$. The (random) set $C_u(x)$ is defined as the set of all points joined to x by open paths not using the edge $\langle 0, u\rangle$.

In the next lemma, we have a strictly positive integer R, and we let $B = B(R)$. The set $C_B(x)$ is defined as the set of all points reachable from x along open paths using no vertex of B.

(10.64) Lemma. *Let u be a (positive or negative) unit vector. We have that*

$$P_p\big(0 \leftrightarrow x,\ u \leftrightarrow y \text{ off } C_u(x)\big) \geq \alpha(p) P_p\big(0 \leftrightarrow x,\ u \leftrightarrow y \text{ off } C_B(x)\big),$$

for $x, y \in \mathbb{Z}^d$, where $\alpha(p) = \{\min(p, 1 - p)\}^{2d(2R+1)^d}$.

Proof. Define the following events,

$$\begin{aligned} E &= \big\{0 \leftrightarrow x,\ u \leftrightarrow y \text{ off } C_u(x)\big\}, \\ F &= \big\{0 \leftrightarrow x,\ u \leftrightarrow y \text{ off } C_B(x)\big\}, \\ G &= \big\{B \cap C(x) \neq \varnothing,\ B \cap C(y) \neq \varnothing,\ C_B(x) \cap C_B(y) = \varnothing\big\}, \end{aligned} \tag{10.65}$$

noting that $E \subseteq F \subseteq G$. The event F is illustrated in Figure 10.3. Now,

$$P_p(E) = P_p(E \mid G) P_p(G) \geq P_p(E \mid G) P_p(F).$$

The event G is independent of the states of all edges lying in the edge-set $\mathbb{E}_B$ of B. Also, for any $\omega \in G$, there exists a configuration $\omega_B = \omega_B(\omega)$ for the edges in $\mathbb{E}_B$ such that the composite configuration (ω off $\mathbb{E}_B$, and ω_B on $\mathbb{E}_B$) lies in E. Since $\mathbb{E}_B$ is finite, and $P_p(\omega_B) \geq \alpha(p)$ whatever the choice of ω_B, we have that $P_p(E \mid G) \geq \alpha(p)$, and the conclusion of the lemma follows. □

For $A \subseteq \mathbb{Z}^d$ and $u, y \in \mathbb{Z}^d$, let $\tau_p^A(u, y) = P_p(u \leftrightarrow y \text{ off } A)$. Now,

$$\begin{aligned} \tau_p(u, y) &= \tau_p^A(u, y) + P_p\big(u \leftrightarrow y, \text{ but } u \not\leftrightarrow y \text{ off } A\big) \\ &\leq \tau_p^A(u, y) + \sum_{w \in A} P_p\big(\{u \leftrightarrow w\} \circ \{y \leftrightarrow w\}\big) \\ &\leq \tau_p^A(u, y) + \sum_{w \in A} \tau_p(u, w) \tau_p(y, w) \end{aligned} \tag{10.66}$$

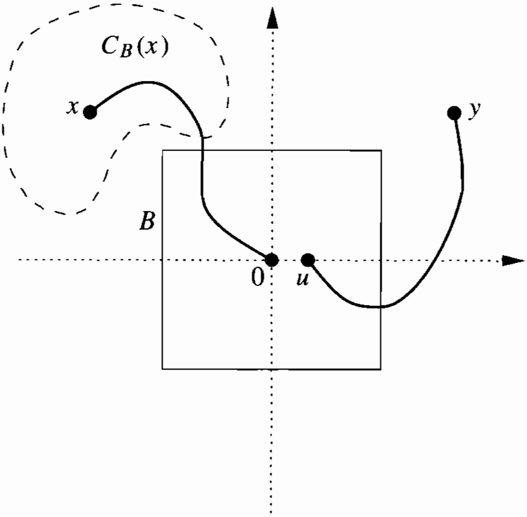


Figure 10.3. An illustration of the event that: $0 \leftrightarrow x$, and $u \leftrightarrow y$ off $C_B(x)$.

by the BK inequality. This equation is valid for all sets A, and we are free to choose A to be a random set.

By (10.63) and Lemma (10.64),

$$\frac{d\chi}{dp} \geq \alpha(p) \sum_{x,y} \sum_{|u|=1} P_p\big(0 \leftrightarrow x,\ u \leftrightarrow y \text{ off } C_B(x)\big). \tag{10.67}$$

Next, we condition on the random set $C_B(x)$. For given $D \subseteq \mathbb{Z}^d$, the event $\{C_B(x) = D\}$ depends only on the states of edges in $\mathbb{Z}^d \setminus B$ having at least one endpoint in D; in particular, we have no information about the states of edges which either touch no vertex of D, or which touch at least one vertex of B. We may therefore apply the FKG inequality to obtain that

$$P_p\big(0 \leftrightarrow x,\ u \leftrightarrow y \text{ off } C_B(x) \,\big|\, C_B(x)\big) \geq P_p\big(0 \leftrightarrow x \,\big|\, C_B(x)\big)\tau_p^{C_B(x)}(u, y).$$

Hence,

$$\begin{aligned} P_p\big(0 \leftrightarrow x,\ u \leftrightarrow y \text{ off } C_B(x)\big) &\geq E_p\Big(E_p\big(I_{\{0\leftrightarrow x\}}\tau_p^{C_B(x)}(u, y) \,\big|\, C_B(x)\big)\Big) \\ &= E_p\big(I_{\{0\leftrightarrow x\}}\tau_p^{C_B(x)}(u, y)\big), \end{aligned}$$

and therefore,

$$\begin{aligned} &P_p\big(0 \leftrightarrow x,\ u \leftrightarrow y \text{ off } C_B(x)\big) \geq \\ &\qquad \tau_p(0, x)\tau_p(u, y) - \Big[E_p\big(I_{\{0\leftrightarrow x\}}\tau_p(u, y)\big) - E_p\big(I_{\{0\leftrightarrow x\}}\tau_p^{C_B(x)}(u, y)\big)\Big]. \end{aligned}$$

Applying (10.66) with $A = C_B(x)$, we have that

$$\begin{aligned}\tau_p(u,y) - \tau_p^{C_B(x)}(u,y) &\le \sum_{w\in C_B(x)} \tau_p(u,w)\tau_p(y,w)\\ &= \sum_{w\in\mathbb{Z}^d\setminus B} I_{\{w\to x \text{ off } B\}}\tau_p(u,w)\tau_p(y,w),\end{aligned}$$

whence

$$(10.68)\quad P_p\big(0\leftrightarrow x,\ u\leftrightarrow y \text{ off } C_B(x)\big) \ge \tau_p(0,x)\tau_p(u,y) - \sum_{w\in\mathbb{Z}^d\setminus B} P_p\big(0\leftrightarrow x,\ w\leftrightarrow x \text{ off } B\big)\tau_p(u,w)\tau_p(y,w).$$

Finally, using the BK inequality as in (6.88),

$$P_p\big(0\leftrightarrow x,\ w\leftrightarrow x \text{ off } B\big) \le \sum_{v\in\mathbb{Z}^d\setminus B} \tau_p(0,v)\tau_p(w,v)\tau_p(x,v).$$

We insert this into (10.68), and deduce via (10.67) that

$$(10.69)\qquad \frac{d\chi}{dp} \ge 2d\alpha(p)\chi^2\left(1 - \sup_{|u|=1}\sum_{v,w\in\mathbb{Z}^d\setminus B} \tau_p(0,v)\tau_p(v,w)\tau_p(w,u)\right).$$

Let

$$Q(u) = \sum_{v,w\in\mathbb{Z}^d} \tau_p(0,v)\tau_p(v,w)\tau_p(w,u)$$

as above (10.60). By (10.59)–(10.60),

$$Q(u) \le Q(0) = T(p),$$

which is to say that

$$(10.70)\qquad \sum_{v,w} \tau_p(0,v)\tau_p(v,w)\tau_p(w,u) \le T(p) \quad \text{for all } u.$$

Assuming that $T(p_c) < \infty$, we may choose $B = B(R)$ sufficiently large that

$$\sum_{v,w\in\mathbb{Z}^d\setminus B} \tau_p(0,v)\tau_p(v,w)\tau_p(w,u) \le \tfrac{1}{2} \quad \text{for } p \le p_c,$$

whence

$$(10.71)\qquad \frac{d\chi}{dp} \ge d\alpha(p)\chi^2 \quad \text{for } p \le p_c.$$

We integrate this, as in (10.39), to obtain that

$$(10.72) \qquad \chi(p) \leq \frac{1}{\alpha'(p_c - p)} \quad \text{for } p \leq p_c$$

where $\alpha' = \alpha'(p)$ is strictly positive and continuous for $0 < p < 1$ (and we have used the fact (6.52) that $\chi(p_c) = \infty$). □

Finally, we discuss the verification of the triangle condition $T(p_c) < \infty$. This has been proved for large d (currently $d \geq 19$) by Hara and Slade (1989a, b, 1990, 1994, 1995), and is believed to hold for $d \geq 7$. The corresponding condition for a 'spread out' percolation model, having long finite-range links rather than nearest-neighbour only, is known to hold for $d > 6$.

The proof that $T(p_c) < \infty$ is long and technical, and is to be found in Hara and Slade (1990). The details are not included here; instead, we survey briefly the structure of the proof. The fundamental estimate is an upper bound for $\widehat{\tau}_p$, namely the so called *infra-red bound*:

$$(10.73) \qquad \widehat{\tau}_p(\theta) \leq \frac{c}{|\theta|^2} \quad \text{for some } c\ (< \infty) \text{ and all } p < p_c,$$

where $|\theta| = \sqrt{\theta \cdot \theta}$. It is immediate via (10.60) that, if $d > 6$, the infra-red bound (10.73) implies $T(p) < \infty$; in addition, by the monotonicity and continuity of τ_p as a function of p,

$$T(p_c) = \lim_{p \uparrow p_c} T(p) < \infty.$$

It is believed that

$$(10.74) \qquad \widehat{\tau}_{p_c}(\theta) \asymp \frac{1}{|\theta|^{2-\eta}} \quad \text{as } \theta \to 0$$

where η is the critical exponent given in Table 9.2.

(10.75) Theorem. Infra-red bound. *There exists D satisfying $D > 6$ such that, if $d \geq D$,*

$$\widehat{\tau}_p(\theta) \leq \frac{c(p)}{|\theta|^2}, \quad \theta \in [-\pi, \pi]^d,$$

for some $c(p)$ which is uniformly bounded for $p < p_c$. Also $T(p_c) < \infty$.

The spirit of the proof of the infra-red bound may be summarized in the following way, although the full picture is more complicated. We establish and utilize the following three facts:

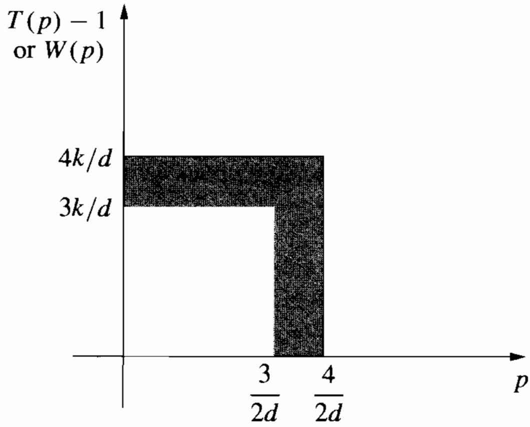


Figure 10.4. There is a 'forbidden region' for the pairs $(p, T(p) - 1)$ and $(p, W(p))$, namely the shaded region in this figure. The quantity k denotes k_T or k_W as appropriate.

(a) the functions $T(p)$ and

$$W(p) = \sum_{x \in \mathbb{Z}^d} |x|^2 \tau_p(0, x)^2$$

are continuous for $p \leq p_c$;

(b) there exist constants k_T and k_W such that

$$T(p) \leq 1 + \frac{k_T}{d}, \quad W(p) \leq \frac{k_W}{d}, \qquad \text{for } p \leq \frac{1}{2d};$$

(c) in addition, for large d, and for p satisfying $(2d)^{-1} \leq p < p_c$, we have that

$$T(p) \leq 1 + \frac{3k_T}{d}, \quad W(p) \leq \frac{3k_W}{d}, \quad p \leq \frac{3}{2d},$$

whenever

$$T(p) \leq 1 + \frac{4k_T}{d}, \quad W(p) \leq \frac{4k_W}{d}, \quad p \leq \frac{4}{2d}.$$

Fact (a) is a consequence of monotone convergence and the continuity of τ_p. Fact (b) follows by comparison with a simpler model (the required comparison is successful for sufficiently small p, namely $p \leq (2d)^{-1}$). Fact (c) is much harder to prove, and it is here that the 'lace expansion' is used. Part (c) implies that there is a 'forbidden region' for the pairs $(p, T(p) - 1)$ and $(p, W(p))$; see Figure 10.4. Since T and W are finite for small p, and continuous up to p_c, part (c) implies that

$$T(p_c) \leq 1 + \frac{3k_T}{d}, \quad W(p_c) \leq \frac{3k_W}{d}, \quad p_c \leq \frac{3}{2d}.$$

The infra-red bound emerges in the proof of (c), of which there follows an extremely brief account.

We write $x \Leftrightarrow y$, and say that x is 'doubly connected' to y, if there exist two edge-disjoint open paths from x to y. We express $\tau_p(0, x)$ in terms of the 'doubly connected' probabilities $\delta_p(u, v) = P_p(u \Leftrightarrow v)$. In doing so, we encounter formulae involving convolutions, which may be treated by taking transforms. At the first stage, we have that

$$\{0 \leftrightarrow x\} = \{0 \Leftrightarrow x\} \cup \left\{ \bigcup_{u \sim v} \{0 \Leftrightarrow (u, v) \leftrightarrow x\} \right\}$$

where $\{0 \Leftrightarrow (u, v) \leftrightarrow x\}$ denotes the event that (i) the edge $\langle u, v \rangle$ is the 'first pivotal edge' for the event $\{0 \leftrightarrow x\}$, (ii) 0 is doubly connected to u, and (iii) the edge $\langle u, v \rangle$ is open. The above union is taken over all ordered pairs u, v of neighbours. (Similar but more complicated events appear throughout the proof.) Therefore,

$$\tau_p(0, x) = \delta_p(0, x) + \sum_{u \sim v} P_p\big(0 \Leftrightarrow (u, v) \leftrightarrow x\big). \tag{10.76}$$

Let $C_{\langle u,v \rangle}(0)$ be the set of vertices reachable from 0 along open paths not using the edge $\langle u, v \rangle$, and let $A(0, u; v, x) = \{v \leftrightarrow x \text{ off } C_{\langle u,v \rangle}(0)\}$. We have that

$$\begin{aligned} P_p&\big(0 \Leftrightarrow (u, v) \leftrightarrow x\big) \\ &= pP_p\big(0 \Leftrightarrow u,\ A(0, u; v, x)\big) \\ &= p\delta_p(0, u)\tau_p(v, x) - pE_p\Big(I_{\{0 \Leftrightarrow u\}}\big\{\tau_p(v, x) - I_{A(0,u;v,x)}\big\}\Big) \end{aligned}$$

whence, by (10.76),

$$\tau_p(0, x) = \delta_p(0, x) + \delta_p \star (pI) \star \tau_p(0, x) - R_{p,0}(0, x) \tag{10.77}$$

where $\star$ denotes convolution, I is the nearest-neighbour function

$$I(u, v) = \begin{cases} 1 & \text{if } u \sim v, \\ 0 & \text{otherwise}, \end{cases}$$

and $R_{p,0}$ is a remainder.

Equation (10.77) is the first step of the lace expansion, In the second step, the remainder $R_{p,0}$ is expanded similarly, and the process is iterated. Such further expansions yield the lace expansion: if $p < p_c$,

$$\tau_p(0, x) = h_{p,N}(0, x) + h_{p,N} \star (pI) \star \tau_p(0, x) + (-1)^{N+1} R_{p,N}(0, x) \tag{10.78}$$

for appropriate remainders $R_{p,N}$, and where

$$h_{p,N}(0, x) = \delta_p(0, x) + \sum_{j=1}^{N} (-1)^j \Pi_{p,j}(0, x)$$

and the $\Pi_{p,j}$ are appropriate functions involving nested expectations of quantities related to 'double connections' (see Hara and Slade (1994), Theorem 4.2).

We take Fourier transforms of (10.78), and solve to obtain that

$$\widehat{\tau}_p = \frac{\widehat{\delta}_p + \sum_{j=1}^{N}(-1)^j \widehat{\Pi}_{p,j} + (-1)^{N+1}\widehat{R}_{p,N}}{1 - p\widehat{I}\widehat{\delta}_p - p\widehat{I}\sum_{j=1}^{N}(-1)^j \widehat{\Pi}_{p,j}}. \tag{10.79}$$

The convergence of the lace expansion, and the consequent validity of this formula for $\widehat{\tau}_p$, is obtained roughly as follows. First, one uses the BK inequality to derive bounds for the δ_p, $\Pi_{p,j}$, $R_{p,j}$ in terms of the functions $T(p)$ and $W(p)$. These bounds then imply bounds for the corresponding transforms. In this way, one may obtain a conclusion which is close to point (c) stated above.

10.4 Notes

Section 10.1. Exact calculations for processes on trees provide a once favoured activity for statistical physicists. Their principal tactics are the standard methods of branching processes, for which the main references are Harris (1963), Athreya and Ney (1972), and Asmussen and Hering (1983). We have borrowed heavily from the treatment presented by Durrett (1985a); see also Section 5 of Kesten (1987b). The idea of studying the asymptotic properties of the moments of the cluster size by way of the random variable Y_p appears to be due to Aizenman and Newman (1984).

Section 10.2. The inequality $\beta \leq 1$ is due to Chayes and Chayes (1987a); see also Menshikov (1987b) and the discussion in the notes for Section 5.2. Aizenman and Newman (1984) proved $\gamma \geq 1$, and Aizenman and Barsky (1987) proved $\delta \geq 2$ and also $\beta \geq (\delta - 1)^{-1}$. The inequality $\beta \geq 2/\delta$ appears in Newman (1987c), along with a similar proof of the inequalities $\gamma, \gamma' \geq 2(1 - \delta^{-1})$, where γ and γ' are the critical exponents for $\chi^{\mathrm{f}}(p)$ when $p < p_{\mathrm{c}}$ and $p > p_{\mathrm{c}}$, respectively; Newman does not assume that $\gamma = \gamma'$. See also Newman (1986, 1987b), and Aizenman, Kesten, and Newman (1987a).

Certain inequalities of debatable interest may be obtained directly from inequality (5.22); assuming only the existence of the critical exponents in question, we may find that $\gamma \leq d$ and $\nu \leq 1$ if $\rho < 1$, whereas $\gamma \leq \rho d$ and $1 \leq \nu \leq \rho$ if $\rho > 1$. See the proof of Theorem (11.93) also.

In further work, Tasaki (1987a, b) has shown that, if the critical exponents for percolation on $\mathbb{L}^d$ equal those for percolation on a binary tree, then $d \geq 6$. Thus the (upper) critical dimension d_{c} cannot be less than 6. Chayes and Chayes (1987b) have reached the same conclusion using different arguments.

Other relevant references are Durrett (1985a), Durrett and Nguyen (1985), Nguyen (1985, 1987a, b), Chayes, Chayes, Fisher, and Spencer (1986), and Newman (1987a).

Rather more is known in the special case when $d = 2$, and we shall discuss such results in the next chapter. Briefly, Kesten (1987b, c) has proved the validity of the scaling and hyperscaling relations

$$2\nu = \gamma + 2\beta = \beta(\delta + 1), \quad 2\rho = \delta + 1, \quad \gamma = \nu(2 - \eta),$$

under the assumption that δ and ρ exist. He has no such exact result for the exponent α, and indeed it remains an open problem to show that $\kappa(p)$ is not thrice differentiable at p_c. Further results concerning hyperscaling may be found in Borgs, Chayes, Kesten, and Spencer (1997). Kesten (1981, 1982) has found bounds for quantities such as θ and χ of the form $\theta(p) \le (p - p_c)^{\sigma}$ for $p > p_c$ and $\chi(p) \le (p_c - p)^{-\tau}$ for $p < p_c$. Such inequalities imply that $\beta > 0$ and $\gamma < \infty$, together with certain other rather basic facts which remain unsettled in three and higher dimensions. Finally, Kesten (1987b) discusses various bounds for critical exponents in two dimensions which imply that many of these exponents cannot equal the corresponding values for percolation on the binary tree. We reproduce his table.

Rigorous bounds for exponents for $d = 2$	'Exact values'	Exponents for binary tree
$\alpha < 0$	$\alpha = -\frac{2}{3}$	$\alpha = -1$
$\beta < 1$	$\beta = \frac{5}{36}$	$\beta = 1$
$\gamma \ge \frac{8}{5}$	$\gamma = \frac{43}{18}$	$\gamma = 1$
$\delta \ge 5$	$\delta = \frac{91}{5}$	$\delta = 2$
$\nu > 1$	$\nu = \frac{4}{3}$	$\nu = \frac{1}{2}$

Table 10.1. Critical exponents for the square lattice and the binary tree.

The so-called 'exact values' are those predicted in the work of den Nijs (1979), Nienhuis, Riedel, and Schick (1980), and Pearson (1980); see also Stauffer (1981), Stauffer and Aharony (1991), and Hughes (1996). We leave this subject now, with the remark that the inequality $\beta < 1$ for two dimensions was proved by Kesten and Zhang (1987). They proved that $\theta(p) \ge A(p - p_c)^b$ for $p > p_c$ and some constants A and $b < 1$; this implies that the right-hand derivative of $\theta(p)$ at p_c is infinite, and justifies our sketch of the function θ in Figure 1.4.

Section 10.3. The triangle condition was introduced by Aizenman and Newman (1984) and used by them to prove (10.54). Barsky and Aizenman (1991) showed that the triangle condition implies (10.53) and more. Nguyen (1987b) deduced (10.56) similarly. Note that (10.53) includes the statement that $\theta(p_c) = 0$.

The proof that the triangle condition holds for large d is due to Hara and Slade (1990, 1994). The method of proof is known as the 'lace expansion', and it may be used directly to prove (10.55); see Hara (1990). Although the validity of the triangle condition is known only for $d \geq 19$, the condition is expected to hold for $d \geq 7$. Hara and Slade have proved this for bond percolation on a different lattice, namely a 'spread out' lattice in which each vertex x is joined to every vertex within some large distance k of x.

Hara and Slade (1995) have used the lace expansion to derive a rigorous expansion for $p_c = p_c(d)$ in terms of powers of d^{-1}, namely equation (3.2). In Hara and Slade (1998), they have discussed material relevant to a construction of the incipient infinite cluster when $p = p_c$ and d is large. They have shown that the Fourier transforms of the two- and three-point connectivity functions

$$P_{p_c}\big(0 \leftrightarrow \lfloor xn^{1/4} \rfloor\big), \; P_{p_c}\big(0 \leftrightarrow \lfloor xn^{1/4} \rfloor \leftrightarrow \lfloor yn^{1/4} \rfloor\big)$$

converge as $n \to \infty$ to natural quantities associated with so called 'integrated super-Brownian excursion'. This work includes the asymptotic formula

$$P_{p_c}(|C| = n) = D\big(1 + O(n^{-\varepsilon})\big)n^{-3/2}$$

where $D, \varepsilon > 0$; hence $\delta = 2$ for large d.

Chapter 11

Bond Percolation in Two Dimensions

11.1 Introduction

Until around 1986, percolation was a game that was played largely on the plane. There is a special reason why percolation in two dimensions is more approachable than percolation in higher dimensions. To every planar two-dimensional lattice $\mathcal{L}$ there corresponds a 'dual' planar lattice $\mathcal{L}_d$ whose edges are in one–one correspondence with the edges of $\mathcal{L}$; furthermore, in a natural embedding of these lattices in the plane, every finite connected subgraph of $\mathcal{L}$ is surrounded by a circuit of $\mathcal{L}_d$. Suppose that we designate a dual edge e_d *open* if and only if the corresponding edge e of the primal is open. This generates a percolation process on the dual lattice $\mathcal{L}_d$, and in this dual pair of processes, the origin of $\mathcal{L}$ belongs to an infinite open cluster if and only if it lies in the interior of no closed circuit of $\mathcal{L}_d$. Such observations may be used to show for some lattices $\mathcal{L}$ that, if $p \neq p_c(\mathcal{L})$, then (almost surely) $\mathcal{L}$ contains an infinite open cluster if and only if the closed clusters of $\mathcal{L}_d$ are finite; it follows that

$$p_c(\mathcal{L}) + p_c(\mathcal{L}_d) = 1, \tag{11.1}$$

where $p_c(\mathcal{L})$ and $p_c(\mathcal{L}_d)$ are the associated critical probabilities. We saw a similar argument in the proof of Theorem (1.10), where it was shown that the square lattice is self-dual in the sense that the dual lattice of $\mathbb{L}^2$ is isomorphic to $\mathbb{L}^2$. Equation (11.1) implies immediately in this case that $p_c(\mathbb{L}^2) = \frac{1}{2}$, the celebrated exact calculation proved by Kesten (1980a) using arguments based on work of Harris, Russo, Seymour, and Welsh.

We explore the consequences of planar duality in this chapter. We shall see that duality provides a method for describing the geometry of open clusters in two dimensions, and consequently the two-dimensional case is better understood

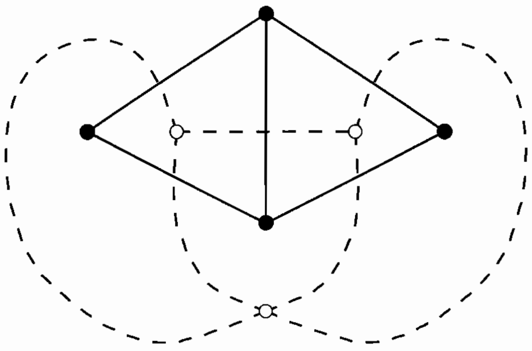


Figure 11.1. A graph and its dual graph.

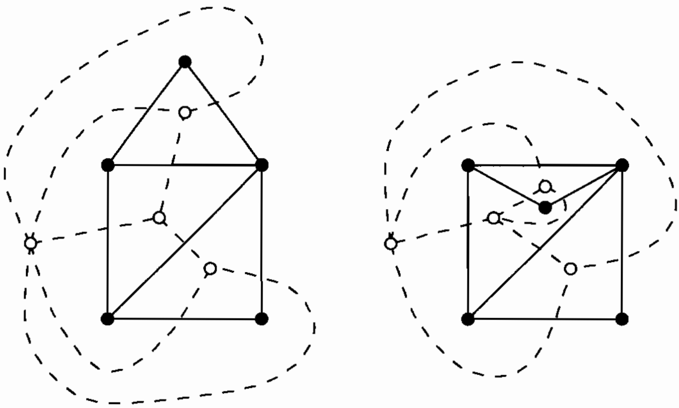


Figure 11.2. Two planar embeddings of the same graph. Note that the duals are not isomorphic, the first containing a vertex having degree five and the second containing only vertices having degrees three and four.

than is its higher-dimensional sibling. Two-dimensional percolation is however interesting not only in its own right, but also because it provides conclusions which, via 'block' arguments, underpin our understanding of all percolation models with $d \geq 2$.

We begin this chapter with a discussion of the technique of planar duality. We shall then apply this technique in proving that $p_c(\mathbb{L}^2) = \frac{1}{2}$. One of the steps of the proof is to show that $\theta(\frac{1}{2}) = 0$, which is to say that there exists (almost surely) no infinite open cluster *at* the critical point. In Section 11.4 we consider the supercritical percolation process, and we show how duality may be used to study both the exponential decay of the truncated connectivity functions, and the distribution of the size of a large finite open cluster. There follows a discussion of percolation on subsets of $\mathbb{L}^2$, and particularly an account of the problem of ascertaining how the critical probability of percolation on the subset S of the

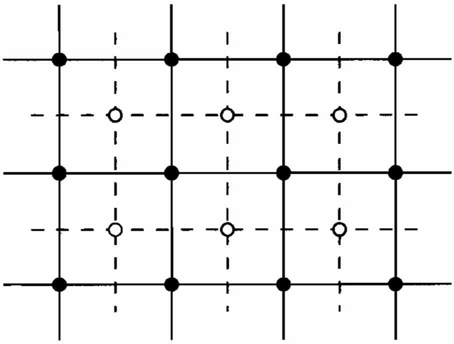


Figure 11.3. Part of the square lattice and its dual. Note that the dual lattice is isomorphic to the original lattice.

lattice depends on the geometry of S. Section 11.6 contains a brief survey of central limit theorems for the contents of large finite regions of the lattice; we deal there with such quantities as the number of open clusters contained within a large contour γ, and the number of vertices inside γ which are joined to vertices of γ by open paths. Section 11.7 is devoted to open crossings of large rectangles, and Section 11.8 to rigorous results for power law behaviour of $\theta(p)$, $\chi(p)$, and related quantities, when p is near to p_c. The chapter terminates with an account of inhomogeneous bond percolation on the square and triangular lattices.

Throughout this chapter, the letter 'd' stands for 'dual' rather than 'dimension' unless the contrary is clear from the context.

11.2 Planar Duality

Let G be a planar graph, drawn in the plane in the manner of Figure 11.1. With G, we may associate another graph G_d called its *(planar) dual*, and we do this in the following way. In each face of G (including the infinite face, if it exists) we place a vertex of G_d; for each edge e of G, we place a corresponding edge joining those two vertices of G_d which lie in the two faces of G abutting e. See Figure 11.1 again for an illustration of this construction. Some care is needed in constructing dual graphs since the dual of a planar graph G depends on the way in which G is embedded in the plane; see Figure 11.2.

We are concerned here with the dual graphs of lattices. A glance at Figure 11.3 indicates that the dual of the square lattice $\mathbb{L}^2$ is isomorphic to $\mathbb{L}^2$; for the sake of being concrete, we shall take the dual of $\mathbb{L}^2$ to be the lattice $\mathbb{L}^2_d$ with vertex set $\{x + (\frac{1}{2}, \frac{1}{2}) : x \in \mathbb{Z}^2\}$ and edges joining all pairs of vertices which are unit distance apart. Thus each edge of $\mathbb{L}^2$ is bisected by an edge of $\mathbb{L}^2_d$ and vice versa. We denote by 0_d the *origin* $(\frac{1}{2}, \frac{1}{2})$ of $\mathbb{L}^2_d$.

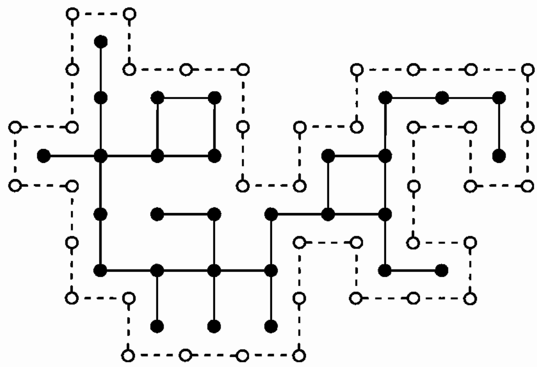


Figure 11.4. A finite connected subgraph of $\mathbb{L}^2$, surrounded by a circuit in the dual each of whose edges traverses an edge in the edge boundary of the subgraph.

Let G be a connected subgraph of $\mathbb{L}^2$, and let ΔG be the edge boundary of G, defined to be the set of edges of $\mathbb{L}^2$ which do not lie in G but which are incident to at least one vertex of G. The following proposition is crucial to the study of percolation on $\mathbb{L}^2$.

(11.2) Proposition. *Let G be a finite connected subgraph of $\mathbb{L}^2$. There exists a unique circuit $\Sigma(G)$ of $\mathbb{L}^2_{\mathrm{d}}$ containing G in its interior and with the property that every edge of $\Sigma(G)$ crosses an edge of ΔG.*

We do not prove this proposition in all its topological glory, but prefer to call it 'obvious' in the light of the picture in Figure 11.4. Kesten (1982, p. 386) presents a more rigorous account of this result. Elementary topological considerations, such as the isoperimetric inequality, may be used as follows to bound the length of the circuit $\Sigma(G)$ referred to in the proposition. First, if G has n vertices then it must be the case that $\Sigma(G)$ has at least $\lambda\sqrt{n}$ vertices, for some absolute positive constant λ; similarly, if $\Sigma(G)$ has m vertices then G must have at least m/ζ vertices for some absolute positive constant ζ. We suppose henceforth that λ and ζ are positive constants such that

$$\lambda\sqrt{|G|} \le |\Sigma(G)| \le \zeta|G| \tag{11.3}$$

for all finite connected subgraphs G of $\mathbb{L}^2$.

Let $0 \le p \le 1$ and consider percolation on $\mathbb{L}^2$ with edge-probability p. We declare each edge of $\mathbb{L}^2_{\mathrm{d}}$ to be *open* (respectively *closed*) if it crosses an open (respectively closed) edge of $\mathbb{L}^2$. This gives rise to a percolation process on $\mathbb{L}^2_{\mathrm{d}}$, which we call the *dual percolation process*. It follows from the self-duality of the square lattice that the statistical properties of the dual process are identical to those of the original process.

Let C be the open cluster of $\mathbb{L}^2$ at the origin. It is a consequence of Proposition (11.2) that C is finite if and only if the origin of $\mathbb{L}^2$ is contained in the interior of a closed circuit of the dual lattice $\mathbb{L}^2_d$, and it was this argument which we used in the proof of Theorem (1.10).

In an early 'exact calculation' of the critical probabilities of certain two-dimensional lattices, Sykes and Essam (1964) presented an argument which is simple, beautiful, and highly plausible. It remains an open problem to make their argument completely rigorous, even though their 'exact conclusions' are now known to be valid (see Theorems (11.115) and (11.116)). Their principal observation is the following theorem, the proof of which is an exercise in the use of duality.

(11.4) Theorem. *For percolation on the square lattice $\mathbb{L}^2$, the number $\kappa(p)$ of open clusters per vertex satisfies*

$$\kappa(p) = \kappa(1-p) + 1 - 2p. \tag{11.5}$$

It is but a small non-rigorous step to deduce from this theorem that $p_c = \frac{1}{2}$, and we argue as follows. Let us suppose that κ is an infinitely differentiable function of p except at p_c, where some derivative of κ fails to exist. A glance at (11.5) indicates that $p_c = 1 - p_c$, and so $p_c = \frac{1}{2}$. A similar argument applied to a dual pair $(\mathcal{L}, \mathcal{L}_d)$ of lattices yields (11.1). Even assuming the fact that $p_c = \frac{1}{2}$, can we justify the hypothesis of this argument? We know from Theorem (6.108) that κ is analytic on $[0, p_c)$, and it follows from (11.5) that κ is analytic on $(1 - p_c, 1]$ also. We shall prove in the next two sections that $p_c = \frac{1}{2}$, so that κ is analytic except possibly at p_c. It is an open problem to show that κ is not infinitely differentiable at p_c. Certainly κ is twice differentiable at p_c (see the discussion after Theorem (11.93)) but it is believed that κ is not thrice differentiable at this point; this is related to the conjecture that the critical exponent α in (9.6) satisfies $-1 \leq \alpha < 0$.

Proof of Theorem (11.4). Of principal value here is Euler's formula (see Bondy and Murty (1976, p. 14) or Wilson (1979, p. 66)). Let G be a finite planar graph, drawn in the plane with $v(G)$ vertices, $e(G)$ edges, $f(G)$ finite faces, and $c(G)$ connected components. Euler's formula states that

$$c(G) = v(G) - e(G) + f(G). \tag{11.6}$$

Suppose that $0 \leq p \leq 1$, and consider percolation on $\mathbb{L}^2$ with edge-probability p. Let $B(n)$ be the box with side-length $2n$ and centre at the origin. Let G_n be the subgraph of $\mathbb{L}^2$ having vertex set $B(n)$ together with all open edges joining such vertices. We apply (11.6) and take expectations to obtain

$$E_p(c(G_n)) = |B(n)| - E_p(e(G_n)) + E_p(f(G_n)). \tag{11.7}$$

We divide by $|B(n)|$ and let $n \to \infty$. From the definition of the function $\kappa(p)$, and particularly Theorem (4.2), we have that

$$\frac{1}{|B(n)|} E_p(c(G_n)) \to \kappa(p) \qquad \text{as } n \to \infty. \tag{11.8}$$

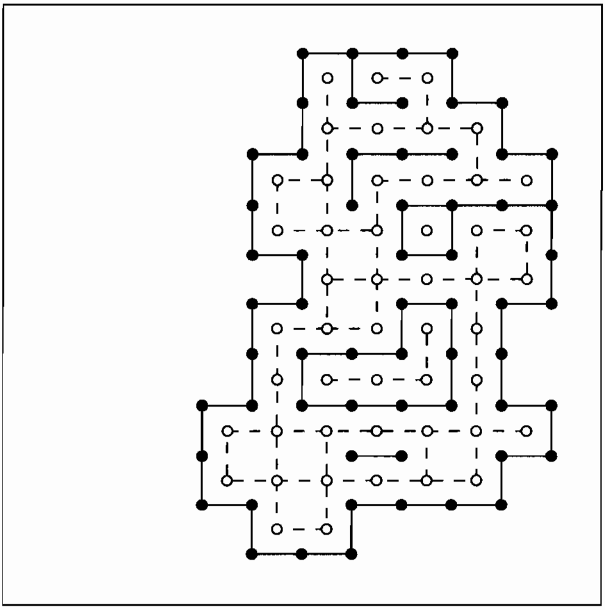


Figure 11.5. The part of G_n sketched here contains three faces, two of which are in the interior of the third. Each such face contains a unique component of the dual.

Also, the mean number of edges in G_n is $2p|B(n)| + o(|B(n)|)$, so that

$$\frac{1}{|B(n)|} E_p(e(G_n)) \to 2p \qquad \text{as } n \to \infty. \tag{11.9}$$

It remains to deal with the mean number of faces of G_n, and it is here that we make use of duality. Let $G_{n,\mathrm{d}}$ be the following subgraph of the dual lattice $\mathbb{L}_d^2$: $G_{n,\mathrm{d}}$ has vertex set $B(n)_\mathrm{d} = \{x + (\frac{1}{2}, \frac{1}{2}) : -n \le x_1, x_2 < n\}$ together with all closed edges joining two such vertices in the dual lattice. We claim that every finite face of G_n contains a unique connected component of $G_{n,\mathrm{d}}$, and we argue as follows. First note that no edge of $G_{n,\mathrm{d}}$ crosses an edge of G_n, since every edge of $G_{n,\mathrm{d}}$ is closed and every edge of G_n is open. Every finite face of G_n contains some vertex of $G_{n,\mathrm{d}}$, and hence some component of $G_{n,\mathrm{d}}$ also. On the other hand, no finite face of G_n can contain two or more such components since, if u and v are vertices of $G_{n,\mathrm{d}}$ which are contained in the same finite face of G_n, then there exists a closed path of $B(n)_\mathrm{d}$ joining u to v. This fact requires more careful proof than that given here; the reader is referred to Figure 11.5 for an illustration of the argument, and to Kesten (1982, p. 244) for more analytical details. Thus each finite face of G_n contains a unique component of $G_{n,\mathrm{d}}$. How many components of $G_{n,\mathrm{d}}$ do not lie in the finite faces of G_n? Any such component lies in the infinite face of G_n and therefore contains a vertex u of $G_{n,\mathrm{d}}$ lying on its boundary $\partial B(n)_\mathrm{d}$. There are $o(|B(n)|)$ such vertices, and therefore $o(|B(n)|)$ such components of

$B(n)_{\mathrm{d}}$. We have shown that the number of finite faces of G_n differs from the number of components of $G_{n,\mathrm{d}}$ by at most $o(|B(n)|)$. Hence,

$$\frac{1}{|B(n)|} E_p(f(G_n)) \to \kappa(1-p) \qquad \text{as } n \to \infty, \tag{11.10}$$

since the mean number of closed components of $G_{n,\mathrm{d}}$ is $(1+o(1))|B(n)_{\mathrm{d}}|\kappa(1-p)$ by Theorem (4.2), and $|B(n)_{\mathrm{d}}|/|B(n)| \to 1$ as $n \to \infty$.

We divide through (11.7) by $|B(n)|$ and let $n \to \infty$ to obtain from (11.8)–(11.10) that

$$\kappa(p) = 1 - 2p + \kappa(1-p)$$

as required. □

11.3 The Critical Probability Equals $\frac{1}{2}$

The principal purpose of this section is to prove the following famous exact calculation.

(11.11) Theorem. *The critical probability of bond percolation on* $\mathbb{L}^2$ *equals* $\frac{1}{2}$.

This result is one of the most important milestones in the history of percolation. Here is some intuitive reasoning to justify the value $\frac{1}{2}$. First, if $\theta(\frac{1}{2}) > 0$ then, when $p = \frac{1}{2}$, there exists (almost surely) both an infinite open cluster in $\mathbb{L}^2$ and an infinite closed cluster in $\mathbb{L}^2_{\mathrm{d}}$. The plane is a somewhat confined space for two non-intersecting infinite clusters, and their co-existence seems unlikely. This is an indication that $\theta(\frac{1}{2}) = 0$, giving that $p_c \geq \frac{1}{2}$. On the other hand, if $p < p_c$ then we may accept a picture of many finite open clusters adrift in an ocean of closed edges of the dual. Presumably this ocean contains an infinite closed cluster, so that $1 - p \geq p_c$; thus $p \leq 1 - p_c$ whenever $p < p_c$, which implies that $p_c \leq \frac{1}{2}$.

The two main ingredients of the streamlined proof which follows are the decay of the tail of $|C|$ when $p < p_c$ (Theorem (5.4)), and the (a.s.) uniqueness of the infinite cluster whenever such a cluster exists (Theorem (8.1)); taken in conjunction with the self-duality of the square lattice, these two ingredients guarantee the result. Approached from this point of view, the result is very natural and its proof rather short and accessible, in contrast to the proof which appeared first in the literature. The latter proof was the crowning achievement of four papers published over a period of 21 years. Harris (1960) proved that $\theta(\frac{1}{2}) = 0$, and therefore $p_c \geq \frac{1}{2}$, using arguments of some geometric complexity and ingenuity; as a byproduct he obtained the uniqueness of the infinite cluster in two dimensions. Attention was drawn to the role of the mean cluster size by Russo (1978) and by Seymour and Welsh (1978) in independent but largely equivalent work. Finally, Kesten (1980a) showed how to build on their arguments to obtain the full result. We shall take a quite different route here, using the uniqueness of the infinite cluster in place of Harris's arguments, and appealing to Theorem (5.4) in place of Kesten's argument and the difficult parts of the Russo–Seymour–Welsh technology.

We divide the theorem into two parts, and shall prove first the following lemma.

(11.12) Lemma. *It is the case that* $\theta(\frac{1}{2}) = 0$, *and therefore the critical probability* p_c *satisfies* $p_c \geq \frac{1}{2}$.

Once we have proved that $p_c = \frac{1}{2}$ then this lemma implies that $\theta(p_c) = 0$, so that there exists (almost surely) no infinite open cluster *at* the critical point. As already noted several times, this implies that θ is continuous on the whole interval $[0, 1]$, a statement which is believed to be valid in all dimensions d (≥ 2). A further consequence of this lemma is the following, which is stated here partly for historical reasons.

(11.13) Lemma. Existence of open circuits above p_c. *Suppose that* p *is such that* $\theta(p) > 0$, *and let* n *be a positive integer. With probability* 1, *there exists an open circuit of* $\mathbb{L}^2$ *containing the box* $B(n)$ *in its interior.*

We recall that $B(n)$ is the box with side-length $2n$ and centre at the origin. By the *interior* of a circuit, we mean the bounded region of $\mathbb{R}^2$ contained strictly within the circuit.

Lemmas (11.12) and (11.13) are essentially equivalent to each other. Harris proved (11.13) directly, and deduced (11.12) together with the uniqueness of the infinite open cluster. On the other hand, it is easy to deduce (11.13) from (11.12), and here is a sketch of the argument; it is not difficult to fill in the gaps. Suppose that $\theta(p) > 0$ but that there exists a positive integer n such that: the probability that $B(n)$ is contained in the interior of an open circuit equals $1 - \delta$ where $\delta > 0$. We have from Lemma (11.12) that $p > \frac{1}{2}$. The box $B(n)$ belongs to the interior of no open circuit if and only if some vertex of the dual lattice $\mathbb{L}_d^2$ lying 'just outside' $B(n)$ belongs to an infinite closed cluster of the dual (see Figure 11.6). This observation is easily checked by applying Proposition (11.2) to the connected subgraph G of $\mathbb{L}_d^2$ containing the 'dual box' $B_d = \{x + (\frac{1}{2}, \frac{1}{2}) : -n-1 \leq x_1, x_2 \leq n\}$, together with all edges of the dual joining pairs of such vertices, and all vertices and closed edges of the dual joined to B_d by closed paths; by the proposition, G is finite if and only if it lies in the interior of some open circuit of the original lattice $\mathbb{L}^2$. Therefore, the probability that there exists an infinite closed cluster in the dual is at least δ, and therefore the density $1 - p$ of closed edges exceeds the critical density p_c. That is to say, $1 - p \geq p_c$, and so $p \leq 1 - p_c$. However, $p_c \geq \frac{1}{2}$ and therefore $p \leq \frac{1}{2}$, a contradiction.

We turn now to the proof of Lemma (11.12). In this proof we shall make use of an elementary device which is of value elsewhere: if $A_1, A_2, \ldots, A_m$ are

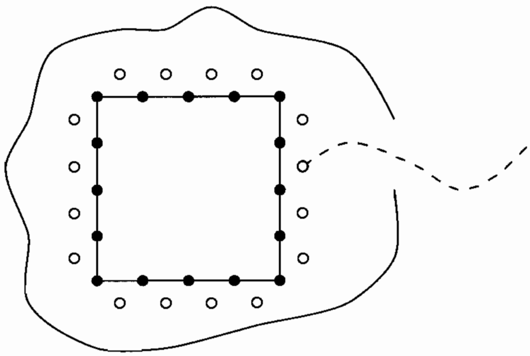


Figure 11.6. If $B(n)$ belongs to the interior of no open circuit, then some vertex of the dual lies in an infinite closed cluster. The broken line represents a closed path in the dual.

increasing events having equal probability, then

$$P_p(A_1) \geq 1 - \left\{1 - P_p\left(\bigcup_{i=1}^m A_i\right)\right\}^{1/m}. \tag{11.14}$$

To see this, note that

$$\begin{aligned} 1 - P_p\left(\bigcup_{i=1}^m A_i\right) &= P_p\left(\bigcap_{i=1}^m \overline{A_i}\right) \\ &\geq \prod_{i=1}^m P_p(\overline{A_i}) \qquad \text{by the FKG inequality} \\ &= \big(1 - P_p(A_1)\big)^m. \end{aligned}$$

We call this the 'square root trick', following Cox and Durrett (1988).

Proof of Lemma (11.12). We follow an argument of Zhang (1988). Consider bond percolation on $\mathbb{L}^2$ with edge-probability $\frac{1}{2}$, and suppose that $\theta(\frac{1}{2}) > 0$. For each positive integer n, let $A^{\mathrm{l}}(n)$ (respectively $A^{\mathrm{r}}(n)$, $A^{\mathrm{t}}(n)$, $A^{\mathrm{b}}(n)$) be the event that some vertex on the left (respectively right, top, bottom) side of the square $T(n) = [0, n]^2$ lies in an infinite open path of $\mathbb{L}^2$ which uses no other vertex of $T(n)$. Clearly $A^{\mathrm{l}}(n)$, $A^{\mathrm{r}}(n)$, $A^{\mathrm{t}}(n)$, and $A^{\mathrm{b}}(n)$ are increasing events having equal probability and whose union is the event that some vertex in $T(n)$ lies in an infinite open cluster. There exists an infinite open cluster with probability 1, whence

$$P_{1/2}\big(A^{\mathrm{l}}(n) \cup A^{\mathrm{r}}(n) \cup A^{\mathrm{t}}(n) \cup A^{\mathrm{b}}(n)\big) \to 1 \qquad \text{as } n \to \infty.$$

It follows by the square root trick (11.14) that

$$P_{1/2}(A^u(n)) \to 1 \qquad \text{as } n \to \infty, \text{ for } u = \mathrm{l}, \mathrm{r}, \mathrm{t}, \mathrm{b}. \tag{11.15}$$

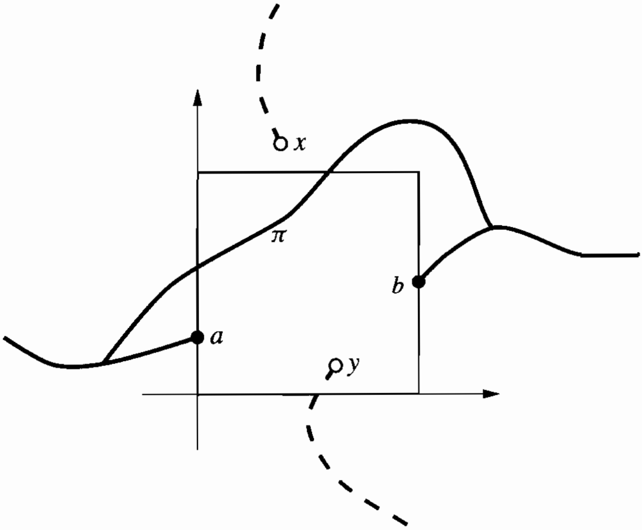


Figure 11.7. Vertices a and b lie in infinite open clusters of $\mathbb{L}^2\backslash T(N)$, and vertices x and y lie in infinite closed clusters of $\mathbb{L}^2_{\mathrm{d}}\backslash T(N)_{\mathrm{d}}$. If there exists a unique infinite open cluster, then there exists an open path π joining a to b, and thus the infinite closed clusters at x and y are disjoint.

We choose N such that

$$(11.16)\qquad P_{1/2}(A^u(N)) > \tfrac{7}{8} \qquad \text{for } u = \mathrm{l}, \mathrm{r}, \mathrm{t}, \mathrm{b}.$$

Moving to the dual lattice, we define the dual box

$$(11.17)\qquad T(n)_{\mathrm{d}} = \left\{x + (\tfrac{1}{2}, \tfrac{1}{2}) : 0 \le x_1, x_2 \le n\right\}.$$

Let $A^{\mathrm{l}}_{\mathrm{d}}(n)$ (respectively $A^{\mathrm{r}}_{\mathrm{d}}(n)$, $A^{\mathrm{t}}_{\mathrm{d}}(n)$, $A^{\mathrm{b}}_{\mathrm{d}}(n)$) be the event that some vertex on the left (respectively right, top, bottom) side of $T(n)_{\mathrm{d}}$ lies in an infinite closed path of $\mathbb{L}^2_{\mathrm{d}}$ which uses no other vertex of $T(n)_{\mathrm{d}}$. Each edge of $\mathbb{L}^2_{\mathrm{d}}$ is closed with probability $\frac{1}{2}$, whence

$$(11.18)\qquad P_{1/2}(A^u_{\mathrm{d}}(N)) = P_{1/2}(A^u(N)) > \tfrac{7}{8} \qquad \text{for } u = \mathrm{l}, \mathrm{r}, \mathrm{t}, \mathrm{b},$$

by (11.16). Consider now the event $A = A^{\mathrm{l}}(N) \cap A^{\mathrm{r}}(N) \cap A^{\mathrm{t}}_{\mathrm{d}}(N) \cap A^{\mathrm{b}}_{\mathrm{d}}(N)$, that there exist infinite open paths of $\mathbb{L}^2\backslash T(N)$ touching the left and right sides of $T(N)$, and also infinite closed paths of $\mathbb{L}^2_{\mathrm{d}}\backslash T(N)_{\mathrm{d}}$ touching the top and bottom sides of $T(N)_{\mathrm{d}}$; see Figure 11.7.

The probability that A does not occur satisfies

$$\begin{aligned} P_{1/2}(\overline{A}) &\le P_{1/2}\big(\overline{A^{\mathrm{l}}(N)}\big) + P_{1/2}\big(\overline{A^{\mathrm{r}}(N)}\big) + P_{1/2}\big(\overline{A^{\mathrm{t}}_{\mathrm{d}}(N)}\big) + P_{1/2}\big(\overline{A^{\mathrm{b}}_{\mathrm{d}}(N)}\big) \\ &\le \tfrac{1}{2} \qquad \text{by (11.16) and (11.18),} \end{aligned}$$

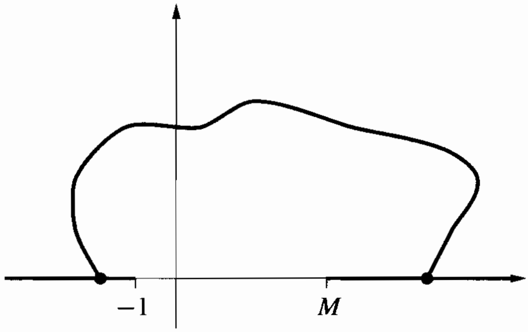


Figure 11.8. A sketch of the event A_M.

giving that $P_{1/2}(A) > \frac{1}{2}$. If A occurs then $\mathbb{L}^2 \backslash T(N)$ contains two disjoint infinite open clusters, since the clusters in questions are separated physically by infinite closed paths of the dual; any open path of $\mathbb{L}^2 \backslash T(N)$ joining these two clusters would contain an edge which crosses a closed edge of the dual, and no such edge can exist. Similarly, on A, the graph $\mathbb{L}^2_{\mathrm{d}} \backslash T(N)_{\mathrm{d}}$ contains two disjoint infinite closed clusters, separated physically by infinite open paths of $\mathbb{L}^2 \backslash T(N)$. Now, the whole lattice $\mathbb{L}^2$ contains (almost surely) a *unique* infinite open cluster, and it follows that there exists (almost surely on A) an open connection of the lattice between the fore-mentioned infinite open clusters. By the geometry of the situation (see Figure 11.7 again), this connection forms a barrier to possible closed connections of the dual joining the two infinite closed clusters. Hence, almost surely on A, the dual lattice contains two or more infinite closed clusters. Since the latter event has probability 0, it follows that $P_{1/2}(A) = 0$ in contradiction of the deduction that $P_{1/2}(A) > \frac{1}{2}$. The initial hypothesis that $\theta(\frac{1}{2}) > 0$ is therefore incorrect, and the proof is complete. □

Proof of Theorem (11.11). It remains only to show that $p_c \le \frac{1}{2}$. There is more than one way of going about this, and we shall give two proofs, both of which originate in the work of Russo (1978) and Seymour and Welsh (1978).

Here is our first proof. We shall show that, if $p < p_c$, there is strictly positive probability that the origin of the dual lattice lies in an infinite closed cluster. Such a cluster exists with positive probability only if the probability $1 - p$, that a given edge is closed, is at least as large as p_c. Thus $1 - p \ge p_c$ whenever $p < p_c$, and this implies that $p_c \le \frac{1}{2}$ as required.

Suppose that $p < p_c$, so that the mean cluster size $\chi(p)$ satisfies

(11.19) $$\chi(p) = \sum_{n=1}^{\infty} P_p(|C| \ge n) < \infty.$$

Let M be a positive integer. We denote by A_M the event that there exists an open path π in $\mathbb{L}^2$ joining a vertex of the form $(k, 0)$ with $k < 0$ to a vertex of the form

$(l, 0)$ with $l \geq M$, and having the property that all the vertices of π other than its endvertices lie strictly above the horizontal axis. A sketch of the event A_M appears in Figure 11.8. We may see that

$$P_p(A_M) \leq P_p\left(\bigcup_{l=M}^{\infty}\big\{(l,0) \leftrightarrow (k,0) \text{ for some } k < 0\big\}\right)$$
$$\leq \sum_{l=M}^{\infty} P_p(|C| \geq l)$$

since, if $(l, 0)$ is joined to some vertex $(k, 0)$ with $k < 0$, then $(l, 0)$ lies in an open cluster of size at least l. By (11.19), we may choose M large enough to ensure that $P_p(A_M) \leq \frac{1}{2}$. We claim that if A_M does not occur then there exists an infinite closed cluster in the dual lattice. Let L be the set of vertices $\{(m + \frac{1}{2}, \frac{1}{2}) : 0 \leq m < M\}$ in the dual, and let $C(L)$ be the set of vertices in the dual which are connected to some vertex in L by a closed path. If $|C(L)| < \infty$, we apply Proposition (11.2) to deduce that $C(L)$ is surrounded in the dual of $\mathbb{L}^2_{\mathrm{d}}$ by an open circuit Σ (this is seen easiest by applying the proposition to the graph with vertex set $C(L)$ and all edges of $\mathbb{L}^2_{\mathrm{d}}$ joining pairs of such vertices). However, the dual of $\mathbb{L}^2_{\mathrm{d}}$ is the original lattice $\mathbb{L}^2$, and such an open circuit Σ must contain an open path joining some vertex $(k, 0)$ with $k < 0$ to some vertex $(l, 0)$ where $l \geq M$, and otherwise lying entirely in the upper half-plane. Thus

$$P_p\big(|C(L)| < \infty\big) \leq P_p(A_M) \leq \tfrac{1}{2},$$

giving that $P_p\big(|C(L)| = \infty\big) > \frac{1}{2}$. Now, if $|C(L)| = \infty$ then one of the M vertices of L lies in an infinite closed cluster of $\mathbb{L}^2_{\mathrm{d}}$, giving that

$$P_p\big(0_{\mathrm{d}} \text{ belongs to an infinite closed cluster}\big) \geq \frac{1}{M} P_p\big(|C(L)| = \infty\big)$$
$$> \frac{1}{2M},$$

and we have shown that $\theta(1 - p) > 0$, as required. The first proof that $p_c \leq \frac{1}{2}$ is now complete.

Our second proof of the inequality $p_c \leq \frac{1}{2}$ makes use of a beautiful application of duality. Let $S(n)$ be the subgraph of $\mathbb{L}^2$ having vertex set $[0, n + 1] \times [0, n]$ and edge set comprising all edges joining pairs of vertices in $S(n)$ *except* those joining pairs x, y with either $x_1 = y_1 = 0$ or $x_1 = y_1 = n + 1$; see Figure 11.9 for a drawing of $S(n)$. Let $S(n)_{\mathrm{d}}$ be the subgraph of $\mathbb{L}^2_{\mathrm{d}}$ having vertex set $\{x + (\frac{1}{2}, \frac{1}{2}) : 0 \leq x_1 \leq n, -1 \leq x_2 \leq n\}$ and edge set comprising all edges of $\mathbb{L}^2_{\mathrm{d}}$ joining pairs of vertices in $S(n)_{\mathrm{d}}$ *except* those joining pairs u, v with either $u_2 = v_2 = -\frac{1}{2}$ or $u_2 = v_2 = n + \frac{1}{2}$. A glance at Figure 11.9 will bring this definition to life; note that $S(n)_{\mathrm{d}}$ may be obtained by rotating $S(n)$

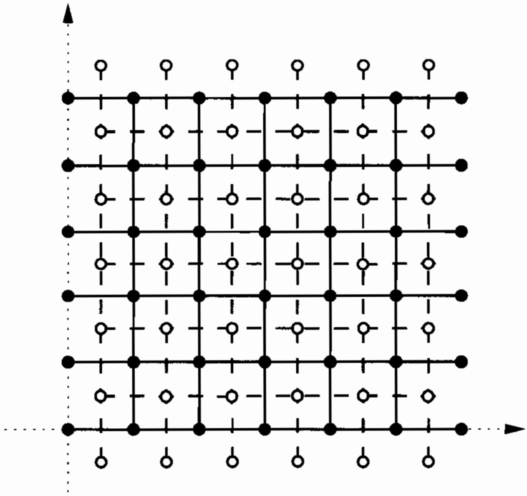


Figure 11.9. A picture of $S(5)$ and its dual $S(5)_{\mathrm{d}}$.

anticlockwise through a right angle, and relocating the vertex labelled (0, 0) at the point $(n+\frac{1}{2}, -\frac{1}{2})$.

We denote by A_n the event that there exists an open path of $S(n)$ joining a vertex on the left side of $S(n)$ to a vertex on its right side. Similarly, we denote by B_n the event that there exists a closed path of $S(n)_{\mathrm{d}}$ joining a vertex on the top side of $S(n)_{\mathrm{d}}$ to a vertex on its bottom side. It is easily seen that $A_n \cap B_n = \varnothing$, the empty event, since if both A_n and B_n occur then there exists an open path in $S(n)$ which crosses a closed path in $S(n)_{\mathrm{d}}$; see Figure 11.9. Where these two paths cross, there is an open edge of $\mathbb{L}^2$ which is crossed by a closed edge of $\mathbb{L}^2_{\mathrm{d}}$, and this is impossible. Thus $A_n \cap B_n = \varnothing$. On the other hand, either A_n or B_n must occur. To see this, suppose that A_n does not occur, and let D be the set of all vertices of $S(n)$ which are attainable from the left side of $S(n)$ along open paths; we turn D into a graph by adding all open edges of $S(n)$ joining pairs of vertices in D. By a minor variant of Proposition (11.2) (illustrated in Figure 11.10) there exists a closed path of $\mathbb{L}^2_{\mathrm{d}}$ crossing $S(n)_{\mathrm{d}}$ from top to bottom, and which crosses only edges of $S(n)$ contained in the edge boundary of D. Thus B_n occurs whenever A_n does not occur. We have proved that A_n and B_n are disjoint events which partition the sample space, and so

$$P_p(A_n) + P_p(B_n) = 1.$$

On the other hand, $P_p(B_n) = P_{1-p}(A_n)$, since $S(n)_{\mathrm{d}}$ is isomorphic to $S(n)$ and each edge of $S(n)_{\mathrm{d}}$ is closed with probability $1-p$. Hence

$$P_p(A_n) + P_{1-p}(A_n) = 1, \tag{11.20}$$

and in particular $P_{1/2}(A_n) = \frac{1}{2}$. We state this conclusion as a lemma, to facilitate its availability for future use.

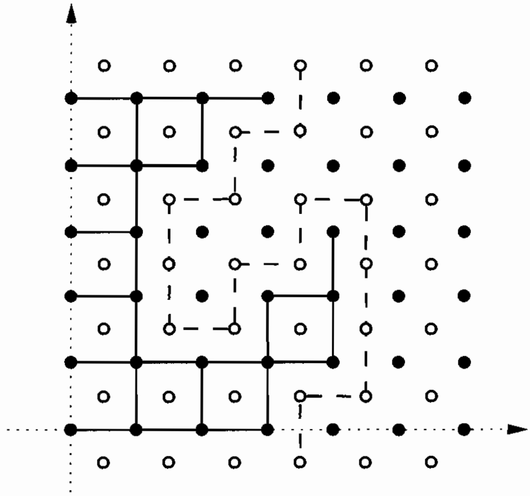


Figure 11.10. If there exists no open path traversing $S(n)$ from left to right, then there exists a closed path crossing $S(n)_{\mathrm{d}}$ from top to bottom.

(11.21) Lemma. *Let A_n be the event that there exists an open path in the rectangle $[0, n+1] \times [0, n]$ joining some vertex on its left side to some vertex on its right side. Then $P_{1/2}(A_n) = \frac{1}{2}$.*

This latter definition of A_n may appear at first sight to differ from the original definition; however, it is easily seen that the two definitions are equivalent, since the existence of such an open crossing does not depend on the states of edges in the left and right sides of the rectangle.

Suppose now that $p_c > \frac{1}{2}$. It follows that the value $p = \frac{1}{2}$ belongs to the subcritical phase, for which the probability of an open path from the origin to a vertex on the line $L_n = \{(n, k) : k \in \mathbb{Z}\}$ decays exponentially as $n \to \infty$. That is to say, if $p_c > \frac{1}{2}$, there exists $\sigma > 0$ such that

$$P_{1/2}(0 \leftrightarrow L_n) \le e^{-\sigma n} \qquad \text{for all } n,$$

by either Theorem (5.4) or Theorem (6.1). In this case,

$$\begin{aligned} P_{1/2}(A_n) &\le \sum_{k=0}^{n} P_{1/2}\big((0,k) \leftrightarrow L_n\big) \\ &\le (n+1)e^{-\sigma n} \\ &\to 0 \qquad \text{as } n \to \infty, \end{aligned}$$

in contradiction of the fact that $P_{1/2}(A_n) = \frac{1}{2}$ for all n. It follows that $p_c \le \frac{1}{2}$, and the second proof is complete. □

Finally we make a note concerning the maximal number M_{n+1} of edge-disjoint open left–right crossings of the box $[0, n+1] \times [0, n]$, thus continuing a discussion begun in Section 2.6. Suppose that $p > \frac{1}{2}$. By (11.20) and the argument at the end of the last proof, we have that

$$P_p(A_n) \geq 1 - (n+1)e^{-\sigma n}$$

for some $\sigma = \sigma(p) > 0$. The following lemma now follows as described after Theorem (2.45).

(11.22) Lemma. *Suppose that $p > \frac{1}{2}$. There exist strictly positive constants $\beta = \beta(p)$, $\gamma = \gamma(p)$ such that*

$$P_p(M_{n+1} \leq \beta n) \leq e^{-\gamma n} \qquad \textit{for all } n \geq 1.$$

11.4 Tail Estimates in the Supercritical Phase

There are two principal results in this direction. The first concerns the probability $\tau_p^{\mathrm{f}}(0, e_n)$ that the origin lies in a finite open cluster containing the vertex $e_n = (n, 0)$, and the second concerns the rate of decay of the cluster size distribution as $n \to \infty$.

In preparation, we recall from Section 8.5 that, for two or more dimensions, the correlation length $\xi(p)$ is defined by the asymptotic relation

$$\tau_p^{\mathrm{f}}(0, e_n) \approx e^{-n/\xi(p)} \qquad \text{as } n \to \infty. \tag{11.23}$$

In the following theorem, we establish a relationship between $\xi(p)$ and $\xi(1-p)$. A similar relationship is unlikely to hold in higher dimensions.

(11.24) Theorem. Exponential decay of the truncated connectivity function. *Suppose that $d = 2$ and $1 > p > p_c$ $(= \frac{1}{2})$. It is the case that $\xi(p) = \frac{1}{2}\xi(1-p)$, and therefore $0 < \xi(p) < \infty$.*

Our second result concerns the asymptotic behaviour of $P_p(|C| = n)$ when $p > p_c$. We saw in Section 8.6 that this probability is at least as large as $\exp(-\gamma(p)n^{1/2})$ for some $\gamma(p) > 0$; furthermore, Theorem (8.65) implies an upper bound with an exponent of order $n^{1/2}$. The situation in two dimensions is much simpler than when $d \geq 3$, and we present next a two-dimensional version of Theorem (8.65) for which the proof is straightforward.

(11.25) Theorem. *Suppose that $d = 2$ and $1 > p > p_c$ ($= \frac{1}{2}$). There exists $\eta(p) > 0$ such that*

$$P_p(|C| = n) \le \exp\big(-\eta(p)n^{1/2}\big) \qquad \textit{for all } n.$$

The existence of the limit

$$\lim_{n\to\infty}\left\{-\frac{1}{\sqrt{n}}\log P_p(|C| = n)\right\}$$

is known; see the notes for this section.

We turn now to the proofs of these theorems. In our proof of Theorem (11.24) we shall need a lemma which is of some interest in its own right. Let k be a positive integer and let T_k be the tube $\{x \in \mathbb{Z}^2 : |x_2| \le k\}$. We write '$0 \leftrightarrow e_n$ in T_k' for the event that there exists an open path of T_k joining the origin to the vertex e_n. It is clear that

(11.26) $$\lim_{k\to\infty} P_p(0 \leftrightarrow e_n \text{ in } T_k) = P_p(0 \leftrightarrow e_n) = \tau_p(0, e_n).$$

(11.27) Lemma. *Suppose that $0 < p < 1$. The limit*

(11.28) $$\varphi_k(p) = \lim_{n\to\infty}\left\{-\frac{1}{n}\log P_p(0 \leftrightarrow e_n \textit{ in } T_k)\right\}$$

exists, for each $k \ge 1$. Furthermore,

(11.29) $$P_p(0 \leftrightarrow e_n \text{ in } T_k) \le \exp\big(-n\varphi_k(p)\big) \quad \textit{for all } n \textit{ and } k.$$

The functions $\varphi_k(p)$ satisfy

(11.30) $$0 < \varphi_k(p) < \infty \qquad \textit{for all } k \ge 1 \textit{ and } 0 < p < 1,$$

(11.31) $$\varphi_k(p) \downarrow \varphi(p) \qquad \textit{as } k \to \infty,$$

where

(11.32) $$\varphi(p) = \lim_{n\to\infty}\left\{-\frac{1}{n}\log \tau_p(0, e_n)\right\}.$$

Thus, for each tube T_k, the two-point connectivity function decays in the manner of an exponential function of the distance between the two points, and furthermore the constant φ_k in the exponent converges to the unrestricted constant φ as the width $2k$ of the tube tends to ∞. If $p < \frac{1}{2}$ then $\varphi(p) = \xi(p)^{-1}$, the reciprocal of the correlation length. See Theorem (6.44) and the discussion around (6.54) for the relevant results in the subcritical phase. It is easy to adapt Lemma (11.27) to obtain a result valid in all dimensions.

Proof of Lemma (11.27). This proof is implicit in Chayes, Chayes, Grimmett, Kesten, and Schonmann (1989) and appears explicitly in Durrett and Schonmann (1988b). It is easy to adapt the proof of Theorem (6.44) to see that the limit in (11.28) exists and that

$$\varphi_k(p) = \inf_n \left\{ -\frac{1}{n} \log P_p(0 \leftrightarrow e_n \text{ in } T_k) \right\}; \tag{11.33}$$

hence (11.29) holds. Certainly $\varphi_k(p) < \infty$ since $P_p(0 \leftrightarrow e_n \text{ in } T_k) \geq p^n$, the probability that each of the n edges in the shortest path from the origin to e_n is open. Here is a direct argument which proves that $\varphi_k(p) > 0$. Let A_i be the event that all edges of T_k joining two vertices of the form (i, r) and $(i+1, r)$ for some r are closed. The A_i are independent events and $P_p(A_i) = (1-p)^{2k+1}$. Also,

$$\begin{aligned} P_p(0 \leftrightarrow e_n \text{ in } T_k) &\leq P_p\big(\overline{A_0} \cap \overline{A_1} \cap \cdots \cap \overline{A_{n-1}}\big) \\ &= \big\{1 - (1-p)^{2k+1}\big\}^n, \end{aligned}$$

giving that

$$\varphi_k(p) \geq -\log\big\{1 - (1-p)^{2k+1}\big\}.$$

Finally we prove (11.31). It is the case that $T_k \subseteq T_{k+1} \subseteq \mathbb{Z}^2$, and so

$$\varphi_k(p) \geq \varphi_{k+1}(p) \geq \varphi(p) \qquad \text{for all } k; \tag{11.34}$$

hence the limit $\lim_{k\to\infty} \varphi_k(p)$ exists and satisfies

$$\lim_{k\to\infty} \varphi_k(p) \geq \varphi(p). \tag{11.35}$$

On the other hand, we combine (11.26) with (11.29) to obtain

$$\tau_p(0, e_n) \leq \lim_{k\to\infty} \exp\big(-n\varphi_k(p)\big) = \exp\left(-n \lim_{k\to\infty} \varphi_k(p)\right).$$

We take logarithms of this inequality, divide by n, let n tend to infinity, and find that

$$\varphi(p) \geq \lim_{k\to\infty} \varphi_k(p)$$

by (11.32). Taken in conjunction with (11.35), this implies the result. □

Proof of Theorem (11.24). We follow Chayes, Chayes, Grimmett, Kesten, and Schonmann (1989). We first derive an upper bound for $\tau_p^{\mathrm{f}}(0, e_n)$, and then we show that this bound is sharp to the 'logarithmic' order.

Suppose that $p_c < p < 1$. If the origin and the vertex $e_n = (n, 0)$ belong to the same finite open cluster, there exists a closed circuit in the dual lattice containing both these vertices in its interior (see Proposition (11.2) and the discussion thereafter). We may see from Figure 11.11 that such a circuit contains two edge-disjoint

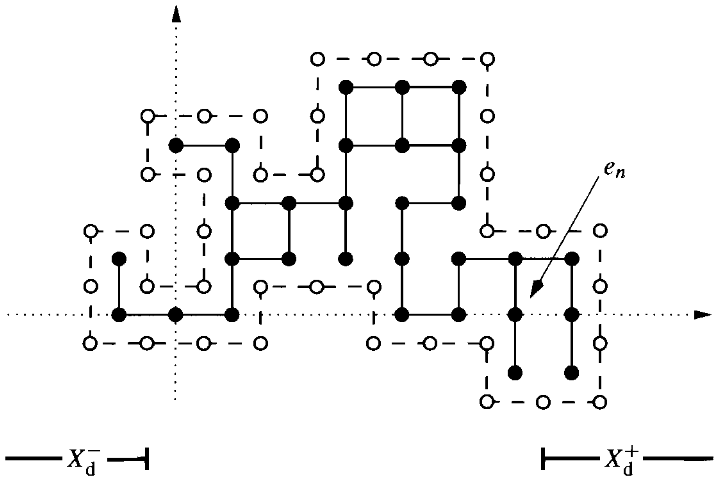


Figure 11.11. Note the existence of two edge-disjoint closed paths joining X_{d}^- to X_{d}^+, each being a sub-path of a closed circuit surrounding the open cluster at the origin.

closed paths joining vertices in X_{d}^- to vertices in X_{d}^+, where X_{d}^- and X_{d}^+ denote the two dual half-axes:

$$X_{\mathrm{d}}^- = \{(k+\tfrac{1}{2}, \tfrac{1}{2}) : k < 0\},$$
$$X_{\mathrm{d}}^+ = \{(l+\tfrac{1}{2}, \tfrac{1}{2}) : l \geq n\}.$$

Writing A_n for the event that some vertex in X_{d}^- is joined by a closed path in the dual lattice to some vertex in X_{d}^+, we have that

$$\tau_p^{\mathrm{f}}(0, e_n) \leq P_p(A_n \circ A_n). \tag{11.36}$$

The event A_n is increasing in the set-theoretic lattice of subsets of the edge-set of $\mathbb{L}_{\mathrm{d}}^2$, and we may therefore apply the BK inequality (particularly inequality (2.17)) to find that

$$\tau_p^{\mathrm{f}}(0, e_n) \leq P_p(A_n)^2. \tag{11.37}$$

Now,

(11.38)

$$\begin{aligned} P_p(A_n) &\leq \sum_{k=-\infty}^{-1} \sum_{l=n}^{\infty} P_p\big((k+\tfrac{1}{2}, \tfrac{1}{2}) \text{ joined to } (l+\tfrac{1}{2}, \tfrac{1}{2}) \text{ by a closed path of the dual}\big) \\ &= \sum_{k=-\infty}^{-1} \sum_{l=n}^{\infty} P_{1-p}\big((k, 0) \leftrightarrow (l, 0)\big). \end{aligned}$$

However, $1 - p < p_c$ since $p > p_c$ $(= \frac{1}{2})$, and so

$$P_{1-p}\big((k,0) \leftrightarrow (l,0)\big) \approx e^{-(l-k)/\xi(1-p)}$$

as $l - k \to \infty$, by (11.23). Furthermore, $\xi(1-p) < \infty$. Therefore,

$$\sum_{k=-\infty}^{-1} \sum_{l=n}^{\infty} P_{1-p}\big((k,0) \leftrightarrow (l,0)\big) \approx e^{-n/\xi(1-p)},$$

and it follows from (11.37) and (11.38) that

$$\text{(11.39)} \qquad \liminf_{n\to\infty} \left\{ -\frac{1}{n} \log \tau_p^{\mathrm{f}}(0, e_n) \right\} \ge \frac{2}{\xi(1-p)},$$

which is one half of the conclusion of the theorem.

For the other half, we need to show that the above approximation to $\tau_p^{\mathrm{f}}(0, e_n)$ is not too bad, and it will follow that

$$\text{(11.40)} \qquad \limsup_{n\to\infty} \left\{ -\frac{1}{n} \log \tau_p^{\mathrm{f}}(0, e_n) \right\} \le \frac{2}{\xi(1-p)};$$

the remainder of the proof is devoted to showing this, and it is here that we make use of Lemma (11.27). Let n be a positive integer and let $m = \lfloor \sqrt{n} \rfloor$; later we shall take the limit as $n \to \infty$. We denote by R_{mn} the rectangle $[-m, n+m] \times [-m, m]$. We write $O(R_{mn})$ for the event that R_{mn} is contained in the interior of some closed circuit of the dual, and $\{0 \leftrightarrow e_n \text{ in } R_{mn}\}$ for the event that there exists an open path in R_{mn} joining the origin to e_n. These two events are defined in terms of disjoint sets of edges and are therefore independent. If they both occur, then the origin and e_n belong to the same finite open cluster of $\mathbb{L}^2$, implying that

$$\text{(11.41)} \qquad \begin{aligned} \tau_p^{\mathrm{f}}(0, e_n) &\ge P_p\big(O(R_{mn}), \text{ and } 0 \leftrightarrow e_n \text{ in } R_{mn}\big) \\ &= P_p(O(R_{mn}))P_p(0 \leftrightarrow e_n \text{ in } R_{mn}). \end{aligned}$$

We need to estimate the last two probabilities, and we begin with $P_p(O(R_{mn}))$. Let k be a positive integer, and define the events

$A_1 = \big\{$there exists a closed path in the dual from $(-m-k-\frac{1}{2}, m+k+\frac{1}{2})$ to $(n+m+k+\frac{1}{2}, m+k+\frac{1}{2})$ lying entirely in the tube $T_1 = \{(u+\frac{1}{2}, v+\frac{1}{2}) : u \in \mathbb{Z},\ m \le v \le m+2k\}\big\}$,

$A_2 = \big\{$there exists a closed path in the dual from $(n+m+k+\frac{1}{2}, m+k+\frac{1}{2})$ to $(n+m+k+\frac{1}{2}, -m-k-\frac{1}{2})$ lying entirely in the tube $T_2 = \{(u+\frac{1}{2}, v+\frac{1}{2}) : n+m \le u \le n+m+2k,\ v \in \mathbb{Z}\}\big\}$,

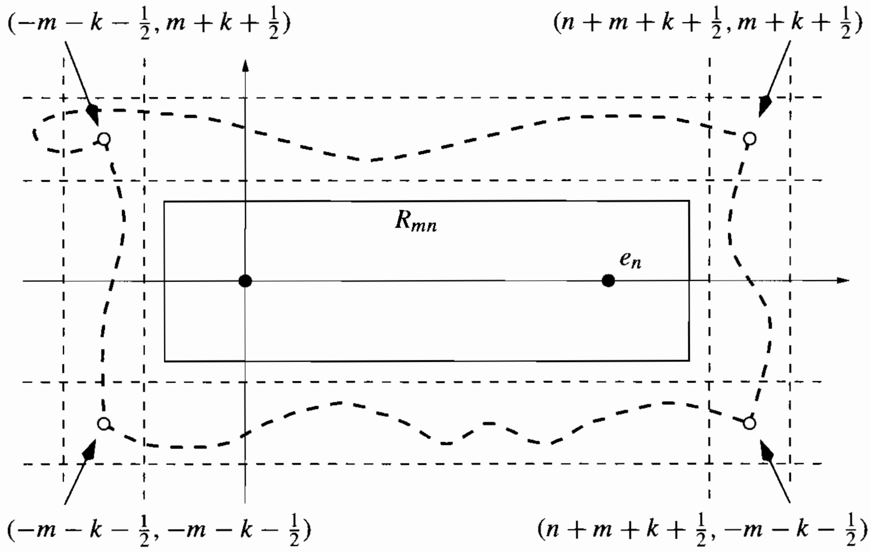


Figure 11.12. The four heavy dashed lines are the four closed paths referred to in the definitions of A_1, A_2, A_3, A_4. If such paths exist then R_{mn} lies in the interior of a closed circuit of the dual.

$$\begin{aligned}
A_3 = \Big\{&\text{there exists a closed path in the dual from}\\
&(n+m+k+\tfrac{1}{2}, -m-k-\tfrac{1}{2}) \text{ to}\\
&(-m-k-\tfrac{1}{2}, -m-k-\tfrac{1}{2}) \text{ lying entirely in the tube}\\
&T_3 = \{(u-\tfrac{1}{2}, v-\tfrac{1}{2}) : u \in \mathbb{Z},\ -m-2k \le v \le -m\}\Big\},
\end{aligned}$$

$$\begin{aligned}
A_4 = \Big\{&\text{there exists a closed path in the dual from}\\
&(-m-k-\tfrac{1}{2}, -m-k-\tfrac{1}{2}) \text{ to}\\
&(-m-k-\tfrac{1}{2}, m+k+\tfrac{1}{2}) \text{ lying entirely in the tube}\\
&T_4 = \{(u-\tfrac{1}{2}, v-\tfrac{1}{2}) : -m-2k \le u \le -m,\ v \in \mathbb{Z}\}\Big\}.
\end{aligned}$$

See Figure 11.12 for a drawing of these events. It is clear that R_{mn} is surrounded by a closed circuit of the dual whenever $A_1 \cap A_2 \cap A_3 \cap A_4$ occurs.

The A_i are decreasing events, and thus we may apply the FKG inequality (and particularly (2.7)) to find that

$$\begin{aligned}
(11.42) \qquad P_p(O(R_{mn})) &\ge P_p(A_1 \cap A_2 \cap A_3 \cap A_4)\\
&\ge P_p(A_1)P_p(A_2)P_p(A_3)P_p(A_4).
\end{aligned}$$

The first and third terms, $P_p(A_1)$ and $P_p(A_3)$, equal the probability that two points, distance $n+2m+2k+1$ apart, are joined by a closed path in a tube of width $2k$; the second and fourth terms equal the probability that two points, distance $2m+2k+1$ apart, are joined by a closed path in a tube of width $2k$. It follows

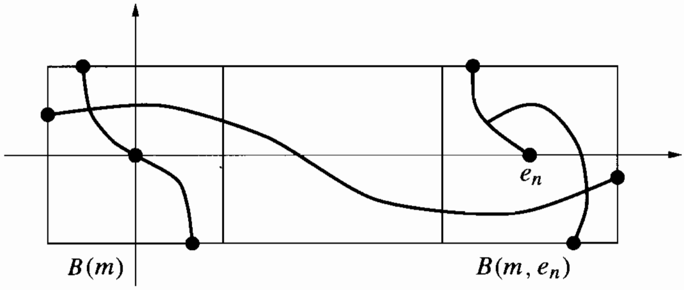


Figure 11.13. Suppose the following three events occur: 0 is joined to the top and bottom sides of $B(m)$, e_n is similarly joined in $B(m, e_n)$, and R_{mn} is traversed from left to right. It follows that 0 is joined to e_n in R_{mn}.

from Lemma (11.27) applied to closed paths that

$$\text{(11.43)} \qquad \frac{1}{n} \log P_p(A_i) \to -\varphi_k(1-p) \qquad \text{as } n \to \infty, \text{ for } i = 1, 3,$$

$$\text{(11.44)} \qquad \frac{1}{n} \log P_p(A_i) \to 0 \qquad \text{as } n \to \infty, \text{ for } i = 2, 4,$$

where φ_k is given in (11.28) and we have used the fact that $m/n \to 0$ as $n \to \infty$. Hence

$$\liminf_{n\to\infty} \left\{ \frac{1}{n} \log P_p(O(R_{mn})) \right\} \geq -2\varphi_k(1-p).$$

This holds for every $k \geq 1$, and so we may take the limit as $k \to \infty$ to obtain

$$\text{(11.45)} \qquad \liminf_{n\to\infty} \left\{ \frac{1}{n} \log P_p(O(R_{mn})) \right\} \geq -2\varphi(1-p) = -\frac{2}{\xi(1-p)},$$

where $\xi(1-p) = \varphi(1-p)^{-1}$ is the correlation length.

We turn next to the last term in (11.41). We write $B(m)$ and $B(m, e_n)$ for the boxes with side-length $2m$ and centres at the origin and e_n, respectively, and we define the following events:

$D_1 = \{$there exists an open path in R_{mn} joining a vertex on its left side to a vertex on its right side$\}$,

$D_2 = \{$there exist open paths in $B(m)$ joining 0 to the upper side and to the lower side of $B(m)\}$,

$D_3 = \{$there exist open paths in $B(m, e_n)$ joining e_n to the upper side and to the lower side of $B(m, e_n)\}$.

See Figure 11.13 for a drawing of these events.

It is clear that there exists an open path in R_{mn} from the origin to e_n whenever $D_1 \cap D_2 \cap D_3$ occurs. The D_i are increasing events, and thus we may use the FKG inequality to find that

$$P_p(0 \leftrightarrow e_n \text{ in } R_{mn}) \geq P_p(D_1)P_p(D_2)P_p(D_3). \tag{11.46}$$

If D_1 does not occur, there exists a closed path in the dual of R_{mn} which separates the left side of R_{mn} from its right side. For this to occur, one of the dual vertices $(r+\frac{1}{2}, m+\frac{1}{2})$, for $-m \leq r < n+m$, must be joined by a closed path in the dual to the line $\{(u+\frac{1}{2}, -m-\frac{1}{2}) : -\infty < u < \infty\}$, and the chance of this is at most

$$\sum_{r=-m}^{n+m-1} P_{1-p}\big(0 \leftrightarrow \partial B(2m)\big) \leq (n+2m)e^{-2m\sigma}$$

for some $\sigma > 0$, by Theorem (6.1). Hence

$$P_p(D_1) \to 1 \qquad \text{as } n \to \infty. \tag{11.47}$$

Finally, $P_p(D_2) = P_p(D_3)$, and so we consider $P_p(D_2)$ only. Let D_4 be the event that there exists an open path of $B(m)$ joining the origin to the upper side of $B(m)$, and let D_5 be the event that a similar path exists to the lower side of $B(m)$. The events D_4 and D_5 are increasing, and their intersection equals D_2. Hence

$$P_p(D_2) = P_p(D_4 \cap D_5) \geq P_p(D_4)P_p(D_5)$$

by the FKG inequality. However, $P_p(D_4) = P_p(D_5)$ by symmetry, and furthermore

$$P_p(D_4) \geq \tfrac{1}{4}\theta(p),$$

since if the origin belongs to an infinite open cluster then it is joined by an open path of $B(m)$ to at least one of the four sides of $B(m)$. We combine these last inequalities to obtain

$$P_p(D_2) = P_p(D_3) \geq \tfrac{1}{16}\theta(p)^2, \tag{11.48}$$

which, considered with (11.46) and (11.47), implies that

$$P_p(0 \leftrightarrow e_n \text{ in } R_{mn}) \geq \big(1+\mathrm{o}(1)\big)\tfrac{1}{16}\theta(p)^2 \qquad \text{as } n \to \infty. \tag{11.49}$$

We return now to (11.41), take logarithms, and use (11.45) and (11.49) to find that

$$\begin{aligned}\limsup_{n\to\infty}\left\{-\frac{1}{n}\log \tau_p^{\mathrm{f}}(0, e_n)\right\} &= -\liminf_{n\to\infty}\left\{\frac{1}{n}\log P_p(O(R_{mn}))\right\} \\ &\leq \frac{2}{\xi(1-p)}\end{aligned}$$

as required in (11.40). The proof is complete. □

Proof of Theorem (11.25). This is a rather straightforward application of the technique of duality. Suppose that $p \in (p_c, 1)$ and that the origin belongs to a finite open cluster of size n. From the remarks after Proposition (11.2), there exists a closed circuit of the dual lattice $\mathbb{L}_d^2$ having the origin of $\mathbb{L}^2$ in its interior and which contains at least $\lambda\sqrt{n}$ vertices, where λ is a positive constant. This circuit must contain some vertex of $\mathbb{L}_d^2$ of the form $(k + \frac{1}{2}, \frac{1}{2})$ where $0 \leq k \leq n - 1$, and therefore some such vertex of $\mathbb{L}_d^2$ must lie in a closed cluster of $\mathbb{L}_d^2$ of size $\lambda\sqrt{n}$ or greater. Each edge of $\mathbb{L}_d^2$ is closed with probability $1 - p$, and $1 - p < p_c$ $(= \frac{1}{2})$ since $p > p_c$ $(= \frac{1}{2})$. Thus the process of closed edges in $\mathbb{L}_d^2$ is subcritical, giving by Theorem (6.78) that there exists $\tau > 0$, depending on p, such that

$$P_p\big(0_d \text{ lies in a closed cluster of } \mathbb{L}_d^2 \text{ with size at least } \lambda\sqrt{n}\big) \leq e^{-\tau\lambda\sqrt{n}}.$$

Therefore,

$$\begin{aligned} P_p(|C| = n) &\leq \sum_{k=0}^{n-1} P_p\big((k + \tfrac{1}{2}, \tfrac{1}{2}) \text{ lies in a closed cluster of } \mathbb{L}_d^2 \text{ with} \\ &\qquad\qquad\qquad\qquad\qquad \text{size at least } \lambda\sqrt{n}\big) \\ &\leq n e^{-\tau\lambda\sqrt{n}}. \end{aligned}$$

We choose $\eta(p)$ accordingly. □

11.5 Percolation on Subsets of the Square Lattice

Let G be a connected subgraph of $\mathbb{L}^2$. The percolation process on $\mathbb{L}^2$ induces a percolation process on G, obtained by deleting all vertices and edges of $\mathbb{L}^2$ which do not lie in G. How do the properties of this induced process depend on the geometry of G? Of particular interest is the question of determining whether or not G contains an infinite open cluster for a given value of p. Clearly, no infinite cluster exists in G when $p \leq \frac{1}{2}$, since the entire lattice contains no such cluster in this case. We may guess that, if $p > \frac{1}{2}$, then all 'reasonably fat' graphs G contain an infinite open cluster. For example, it is not difficult to show that every non-trivial cone contains (almost surely) an infinite cluster whenever $p > \frac{1}{2}$. How 'thin' need the graph G be in order that it contain no infinite open cluster? There is a rather precise answer to this question.

We begin with some notation. Let G be a subset of $\mathbb{Z}^2$; we shall think of G as being a graph whose edges are those of $\mathbb{L}^2$ joining pairs of vertices in G. We shall be particularly interested in sets G of the form

(11.50) $$G(f) = \big\{x \in \mathbb{Z}^2 : 0 \leq x_2 \leq f(x_1),\ x_1 \geq 0\big\},$$

where f is a specified function on $[0, \infty)$ taking non-negative values.

Let p satisfy $0 \le p \le 1$, and consider percolation on $\mathbb{L}^2$ with edge-probability p. We write $\psi(p, G)$ for the probability that there exists an infinite open connected subgraph of G, and we say in this case that 'G contains an infinite open cluster'. As in the case of the lattice $\mathbb{L}^2$ (see Theorem (1.11)), $\psi(p, G)$ is the probability of a tail event of a family of independent random variables, so that $\psi(p, G)$ equals either 0 or 1 for all p and G. We denote by $p_c(G)$ the *critical probability* of G:

$$(11.51) \qquad p_c(G) = \sup\{p : \psi(p, G) = 0\}.$$

An alternative definition of the critical probability $p_c(G)$ is by way of the probability that a given vertex of G belongs to an infinite open cluster of G. Let x be a vertex of G, and let $\theta(p, x; G)$ be the probability that x lies in an infinite open cluster of G. It is not difficult to see as in Theorem (2.8) that, for all pairs x, y of vertices which belong to the same component of G, it is the case that $\theta(p, x; G) > 0$ if and only if $\theta(p, y; G) > 0$. If G is connected, we may define $p_c(G)$ by

$$(11.52) \qquad p_c(G) = \sup\{p : \theta(p, x; G) = 0\},$$

a definition which is independent of the choice of x. By the argument of Theorem (1.11), the two definitions (11.51) and (11.52) of $p_c(G)$ are equivalent whenever G is connected. We shall be concerned here with connected graphs G only, and so we shall use these definitions interchangeably.

Before stating the main result, we recall from Section 6.2 the definition of the correlation length $\xi(p)$ for $p < p_c$ $(= \frac{1}{2})$. Writing $B(n)$ for the usual box having side-length $2n$ and centre at the origin, we have that

$$(11.53) \qquad P_p\big(0 \leftrightarrow \partial B(n)\big) \approx e^{-n/\xi(p)} \qquad \text{as } n \to \infty, \text{ if } 0 < p < \tfrac{1}{2},$$

where ξ is a strictly increasing and continuous function on $(0, \frac{1}{2})$ satisfying $\xi(p) \to 0$ as $p \downarrow 0$ and $\xi(p) \to \infty$ as $p \uparrow \frac{1}{2}$. Furthermore, the probability that there exists an open path from the origin to the vertex $e_n = (n, 0)$ satisfies

$$(11.54) \qquad \tau_p(0, e_n) \approx e^{-n/\xi(p)} \qquad \text{as } n \to \infty.$$

See Figure 11.14 for a sketch of the function ξ.

(11.55) Theorem. *Let $a > 0$, and let f be a non-negative function on $[0, \infty)$ satisfying*

$$\frac{f(u)}{\log u} \to a \qquad \textit{as } u \to \infty.$$

The critical probability of $G(f)$ is the unique number $\pi \in (\frac{1}{2}, 1)$ satisfying

$$\xi(1 - \pi) = a.$$

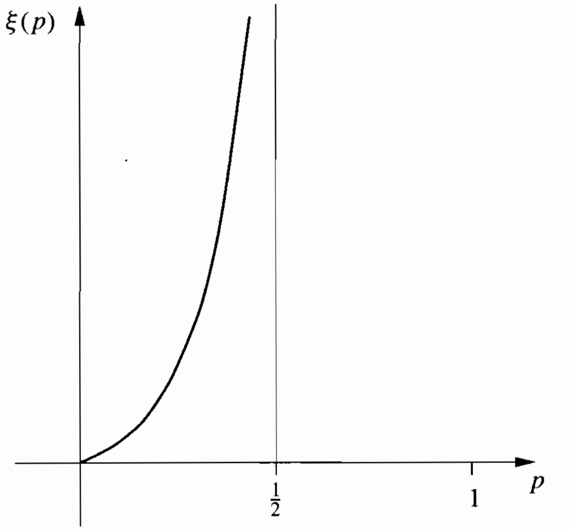


Figure 11.14. A sketch of the function $\xi(p)$ for $p < \frac{1}{2}$.

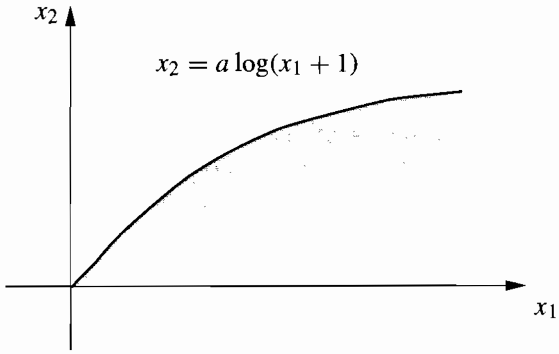


Figure 11.15. A sketch of the region G_a.

We note that ξ maps $(0, \frac{1}{2})$ continuously onto $(0, \infty)$; the strict monotonicity of ξ ensures that the equation $\xi(1 - \pi) = a$ is satisfied by a unique π in $(\frac{1}{2}, 1)$, for any given $a > 0$. The assertion of the theorem amounts to saying that the existence or not of an infinite open cluster in the set

(11.56) $$G_a = \{x \in \mathbb{Z}^2 : 0 \le x_2 \le a \log(x_1 + 1),\ x_1 \ge 0\}$$

depends on whether a is larger or smaller than $\xi(1 - p)$; see Figure 11.15 for a picture of the set G_a.

We may choose the function f in such a way that there exists an infinite open cluster when p *equals* the critical probability for $G(f)$. That is to say, there exist

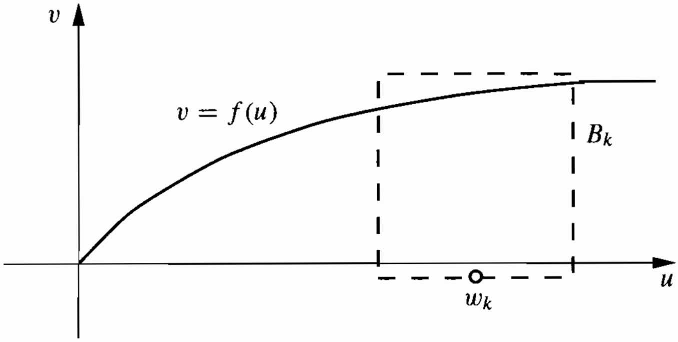


Figure 11.16. The box B_k is the smallest box in the dual having w_k in the middle of its lower side, and its upper side strictly above the curve $v = f(u)$.

functions f such that

$$P_{p_c(G(f))}\big(G(f) \text{ contains an infinite open cluster}\big) = 1.$$

It is not difficult to construct such functions, and we shall see a recipe at the end of the proof of Theorem (11.55): it suffices to let $a > 0$ and to define f by

$$f(u) = a\log(u+1) + b\log\{1 + \log(u+1)\}$$

where $b > 2a$. In the language of mathematical physicists, the corresponding graph $G(f)$ has a 'first order' phase transition in the sense that the probability $\theta(p, x; G(f))$, that the vertex x lies in an infinite open cluster of $G(f)$, is discontinuous at the critical point $p_c(G(f))$.

Proof of Theorem (11.55). We stay close to Grimmett (1981a, 1983). We shall sometimes use real-valued quantities in places where integers are required; this is for notational convenience only and has no mathematical significance. Let $a > 0$, and suppose that f is a function on $[0, \infty)$ which takes non-negative values and satisfies $f(u)/\log u \to a$ as $u \to \infty$.

We consider percolation on $G(f)$, and we suppose first that the edge-probability p satisfies $p > \frac{1}{2}$ and $a < \xi(1-p)$. We shall show that $G(f)$ contains (almost surely) no infinite open cluster. We choose $\delta > 0$ such that $(1+\delta)a < \xi(1-p)$, and we define w_k to be the vertex $(k^{1+\delta} + \frac{1}{2}, -\frac{1}{2})$ of the dual lattice $\mathbb{L}^2_{\mathrm{d}}$. For each k, we construct a box B_k of the dual in the following way. Amongst all square boxes in the dual lattice with w_k in the middle of the lower side, B_k is the smallest with the property that the upper side lies strictly above the curve $v = f(u)$; see Figure 11.16. Now, $f(u)/\log u \to a$ as $u \to \infty$, and it is a simple matter to deduce that the side-length l_k of B_k satisfies

(11.57) $$l_k = a\big(1 + \mathrm{o}(1)\big)\log k^{1+\delta} \qquad \text{as } k \to \infty.$$

We think of B_k as a box with side-length l_k and centre at $w_k + (0, \frac{1}{2}l_k)$. Let A_k be the event that B_k contains a closed path from some vertex on its upper side to some vertex on its lower side. The probability of A_k is no smaller than the probability that the centre of B_k is joined by closed paths of B_k to both its upper and lower sides. Thus

(11.58) $$\begin{aligned} P_p(A_k) &\geq P_p\big(0 \text{ is joined by closed paths to top and bottom sides of } B(\tfrac{1}{2}l_k)\big) \\ &\geq \big\{\tfrac{1}{4}P_{1-p}\big(0 \leftrightarrow \partial B(\tfrac{1}{2}l_k)\big)\big\}^2 \end{aligned}$$

by the FKG inequality, since the probability that the origin is joined to any given side of $B(\frac{1}{2}l_k)$ is at least one quarter of the probability that it is joined to the surface of $B(\frac{1}{2}l_k)$. However,

$$P_{1-p}\big(0 \leftrightarrow \partial B(\tfrac{1}{2}l_k)\big) \approx e^{-l_k(2\xi(1-p))^{-1}} \qquad \text{as } k \to \infty$$

by (11.53). We combine (11.58) with (11.57) and deduce that

$$\begin{aligned} P_p(A_k) &\geq \tfrac{1}{16}\exp\big\{-\big(1+\mathrm{o}(1)\big)l_k/\xi(1-p)\big\} \\ &= \tfrac{1}{16}k^{-(1+\mathrm{o}(1))(1+\delta)a/\xi(1-p)} \end{aligned}$$

as $k \to \infty$, which implies that

(11.59) $$\sum_k P_p(A_k) = \infty$$

since $(1+\delta)a < \xi(1-p)$. On the other hand, the distance from w_k to w_{k+1} is $(k+1)^{1+\delta} - k^{1+\delta} \sim (1+\delta)k^\delta$, and the side-length of B_k is $\mathrm{O}(\log k)$, whence there exists K such that $\{B_k : k \geq K\}$ are disjoint boxes; this implies that $\{A_k : k \geq K\}$ are independent events. Equation (11.59) implies that (almost surely) infinitely many of the events $\{A_k : k \geq K\}$ occur, which implies in turn that $G(f)$ is traversed from top to bottom in the dual by infinitely many closed paths. These closed paths partition $G(f)$ into finite regions (see Figure 11.17). Each open cluster of $G(f)$ is contained within one of these regions, since otherwise there exists an open edge of $\mathbb{L}^2$ crossing a closed edge of $\mathbb{L}^2_{\mathrm{d}}$, and this is impossible. Thus all open clusters of $G(f)$ are (almost surely) finite, and the first part of the proof of complete.

Secondly, we suppose that $p > \frac{1}{2}$ and $a > \xi(1-p)$. Let f be a non-negative function satisfying $f(u)/\log u \to a$ as $u \to \infty$. We shall show that $G(f)$ contains (almost surely) an infinite open cluster. First, we choose α such that $a > \alpha > \xi(1-p)$ and we let D_k be the box of the dual lattice with centre $(k+\frac{1}{2}, \frac{1}{2})$ and side-length $2\alpha \log k$. For all large values of k (say $k \geq L$) it is the case that D_k lies strictly beneath the curve $v = f(u) - 2$. Let E_k be the event that

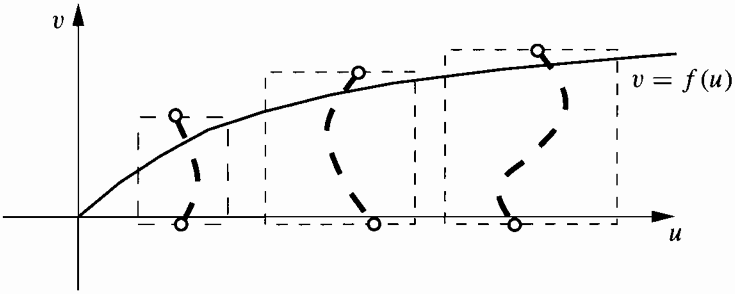


Figure 11.17. The existence of infinitely many closed paths traversing the boxes $\{B_k : k \geq K\}$ prevents the formation of an infinite open cluster in $G(f)$.

$(k + \frac{1}{2}, \frac{1}{2})$ is joined by a closed path of the dual to a vertex on the surface ∂D_k of D_k. We have from (11.53) that

$$\text{(11.60)} \qquad \begin{aligned} P_p(E_k) &= P_{1-p}\big(0 \leftrightarrow \partial B(\alpha \log k)\big) \\ &\approx k^{-\alpha/\xi(1-p)} \qquad \text{as } k \to \infty, \end{aligned}$$

giving that

$$\text{(11.61)} \qquad \sum_k P_p(E_k) < \infty$$

from the assumption $\alpha > \xi(1-p)$. Therefore, there exists $M \geq L$ such that

$$P_p\left(\bigcup_{k \geq M} E_k\right) < \tfrac{1}{2}.$$

However, if none of the events $\{E_k : k \geq M\}$ occurs then no vertex $(k + \frac{1}{2}, \frac{1}{2})$ with $k \geq M$ is joined by a closed path to a vertex of the dual lying above the curve $v = f(u) - 2$. We move to the dual $\mathbb{L}^2$ of $\mathbb{L}^2_{\mathrm{d}}$ to see that there is probability at least $\frac{1}{2}$ that $G(f)$ contains an infinite open path. The probability of the last event is either 0 or 1, and we have proved that $G(f)$ contains almost surely an infinite open cluster, as required.

Finally we show that, for any $a > 0$, there exists a function f such that $f(u)/\log u \to a$ as $u \to \infty$ and

$$P_{p_c(G(f))}\big(G(f) \text{ contains an infinite open cluster}\big) = 1.$$

We consider a function f satisfying

$$f(u) = a \log u + b \log\log u \qquad \text{for all large } u,$$

where $b > 2a$. This time we let D_k be the largest box of the dual lattice having centre at $(k + \frac{1}{2}, \frac{1}{2})$ and lying strictly beneath the curve $v = f(u) - 2$. It is not difficult to show that D_k has side-length $2f(k) + O(1)$ as $k \to \infty$, implying that (11.60) may be replaced by

$$\begin{aligned} P_p(E_k) &= P_{1-p}\big(0 \leftrightarrow \partial B(f(k) + O(1))\big) \\ &\leq \mu(\log k) \exp\left(-\frac{a \log k + b \log\log k}{\xi(1-p)}\right) \end{aligned}$$

for all large k and some absolute constant μ, where we have used the right-hand inequality in (6.11). The critical probability of percolation on $G(f)$ is the number $\pi \in (\frac{1}{2}, 1)$ satisfying $\xi(1 - \pi) = a$, and with this choice of π we have that

$$P_\pi(E_k) \leq \frac{\mu}{k(\log k)^\alpha},$$

where $\pi = p_c(G(f))$ and $\alpha = ba^{-1} - 1 > 1$. It follows as before that

$$P_\pi\big(G(f) \text{ contains an infinite open cluster}\big) = 1$$

as required. □

11.6 Central Limit Theorems

Let K_n be the number of open clusters in the box $B(n)$ when each edge is open with probability p. It was proved in Theorem (4.2) that

$$\frac{1}{|B(n)|} K_n \to \kappa(p) \quad \text{a.s.}$$

as $n \to \infty$, where $\kappa(p) = E_p(|C|^{-1})$ is the number of open clusters per vertex. We may view this limit theorem as a 'strong law' for the number of open clusters. In studying the asymptotic behaviour of sequences of random variables, the classical theory turns next to central limit theorems, and thus we ask whether the sequence $(K_n : n \geq 1)$ is such that

$$\frac{K_n - E_p(K_n)}{\sqrt{\operatorname{var}_p(K_n)}}$$

is asymptotically normally distributed as $n \to \infty$, where var_p denotes variance relative to P_p. This is indeed the case whenever $p \neq 0, \frac{1}{2}, 1$, and this fact is one of a class of central limit theorems for quantities arising out of percolation on $\mathbb{L}^2$.

The general framework is as follows. Any circuit c of the lattice $\mathbb{L}^2$ has an interior and an exterior in the plane $\mathbb{R}^2$. We denote by $\Lambda(c)$ the subgraph of $\mathbb{L}^2$

containing all vertices and edges which are in either c or its interior. We write as usual $\partial\Lambda(c)$ for the surface of $\Lambda(c)$, being the set of vertices in $\Lambda(c)$ which have neighbours which do not belong to $\Lambda(c)$. Suppose that $\mathbf{c} = (c_n : n \geq 1)$ is a sequence of circuits satisfying $|\Lambda(c_n)| \to \infty$ as $n \to \infty$ and that, for each n, we are provided with a function f_n mapping the set of connected subgraphs of $\mathbb{L}^2$ into the real line $\mathbb{R}$. Suppose that $0 < p < 1$, and consider bond percolation on $\mathbb{L}^2$ with edge-probability p. We define the sequence $Z = (Z_n : n \geq 1)$ of random variables by

$$Z_n = \sum_{x \in \Lambda(c_n)} f_n(C(x)), \tag{11.62}$$

where the sum is over all vertices in $\Lambda(c_n)$. We say that 'Z satisfies the central limit theorem' if

$$P_p\left(\frac{Z_n - E_p(Z_n)}{\sqrt{\mathrm{var}_p(Z_n)}} \leq u\right) \to \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} e^{-v^2/2}\, dv$$

for all u, as $n \to \infty$.

This general framework contains many specific examples of interest. For example, let us set c_n to be the bounding circuit of the box $B(n)$, and define

$$f_n(C(x)) = \frac{1}{|C_n(x)|} \qquad \text{for } x \in B(n),$$

where $C_n(x)$ is the set of vertices of $B(n)$ attainable from x along open paths of $B(n)$; then

$$Z_n = \sum_{x \in B(n)} \frac{1}{|C_n(x)|}$$

equals the number K_n of open clusters of $B(n)$, as in Section 4.1.

The aim of this section is to survey known central limit theorems for sequences of the form (11.62), and particularly to discuss the available techniques and their fields of application. We shall give no proofs, but refer the reader to the original papers.

Four basic methods have been employed for proving the asymptotic normality of random variables of the form $Y_n = (Z_n - E_p(Z_n))(\mathrm{var}_p(Z_n))^{-1/2}$.

(1) *Method of moments.* We estimate the semi-invariants (otherwise known as the cumulants) of Y_n and show that they converge to those of the normal distribution. (See Malyshev (1975), Cox and Grimmett (1981, 1984), and the references therein.)

(2) *Association.* If the f_n are increasing functions on the set-theoretic lattice of subgraphs of $\mathbb{L}^2$ (that is to say, $f_n(G) \leq f_n(H)$ whenever G is a subgraph of H), the random variables $(f_n(C(x)) : x \in \Lambda(c_n))$ are positively correlated: this fact can be used to estimate the characteristic function of Y_n directly,

thereby to study its asymptotic behaviour as $n \to \infty$. (See Newman (1980), Newman and Wright (1981), Cox and Grimmett (1984), and the references therein.)

(3) *Strong mixing*. We estimate the correlation coefficients of events related to sets A and B of vertices, as a function of the distance between A and B. This coefficient decays sufficiently rapidly to allow a central limit theorem. (See Neaderhouser (1978, 1981), Grimmett (1979), Bolthausen (1982), Herrndorf (1985), and the references therein.)

(4) *Martingale method*. The sum Z_n may be expressed as a sum of martingale differences. Subject to conditions on these summands, the central limit theorem for martingales may be applied. (See Zhang (1998) and McLeish (1974).)

Before discussing the relative strengths and weaknesses of these three approaches, we state a central limit theorem. We say that a function f on the set of connected subgraphs of $\mathbb{L}^2$ is *constant on infinite graphs* if $f(G) = f(H)$ for all infinite connected subgraphs G and H.

(11.63) Theorem. Central limit theorem. *Suppose that* $\mathbf{c} = (c_n : n \geq 1)$ *is a sequence of circuits of* $\mathbb{L}^2$ *satisfying* $|\Lambda(c_n)| \to \infty$ *as* $n \to \infty$. *Let* $(f_n : n \geq 1)$ *be a sequence of real-valued functions on the set of connected subgraphs of* $\mathbb{L}^2$ *which intersect* $\Lambda(c_n)$. *Suppose that* $0 < p < 1$, *and that*

(11.64) *for all* n, f_n *is a bounded function,*

(11.65) *if* $p > \frac{1}{2}$ *then, for all* n, f_n *is constant on infinite graphs,*

$$\liminf_{n\to\infty} \left\{ \frac{1}{|\Lambda(c_n)|} \operatorname{var}_p(Z_n) \right\} > 0, \tag{11.66}$$

where

$$Z_n = \sum_{x \in \Lambda(c_n)} f_n(C(x)). \tag{11.67}$$

Then $Z = (Z_n : n \geq 1)$ *satisfies the central limit theorem so long as* $p \neq \frac{1}{2}$.

The conditions of this theorem are not the best possible, but this formulation is sufficient for most of our purposes. The result may be proved either by the method of moments (Cox and Grimmett (1981, 1984)) or by the theory of strong mixing (Herrndorf (1985)). It is easy to check whether or not the functions f_n are constant on infinite graphs and bounded. Condition (11.66) can be more difficult to verify, although it is satisfied automatically whenever the f_n are either all increasing or all decreasing and there exists $\sigma^2 > 0$ such that

$$\operatorname{var}_p\big(f_n(C(x))\big) \geq \sigma^2$$

for all $x \in \Lambda(c_n)$ and all n. To see this, just note that

$$\operatorname{var}_p(Z_n) \geq \sum_{x \in \Lambda(c_n)} \operatorname{var}_p\big(f_n(C(x))\big)$$

by the FKG inequality, whenever the f_n are either all increasing or all decreasing.

The condition that $p \neq p_c \ (= \frac{1}{2})$ is essential for the proof of the theorem, since we rely upon the rapid decay of correlations over large distances. The condition of constantness on infinite graphs is important only when $p > \frac{1}{2}$, since if $p < \frac{1}{2}$ then all open clusters are finite with probability 1.

Each of the four techniques of proof has its limitations as well as its advantages, and we discuss some of these next.

(1) *Method of moments.* This is a flexible technique which may be applied in situations where the conditions of a general theorem such as Theorem (11.63) are not satisfied. On the other hand, it seems to require that, for all k, the kth absolute moment of $f_n(C(x))$ is bounded uniformly in n; this has not proved to be an important restriction in practice, since all interesting functions studied so far have been bounded.

(2) *Association.* This is a beautiful and simple argument which seems to be best suited to the case when the regions $\Lambda(c_n)$ are rectangular boxes.

(3) *Strong mixing.* Using such techniques, we reduce to a minimum the moment conditions required of the $f_n(C(x))$. On the other hand, these methods appear at first sight to require that the functions f_n be constant on infinite graphs; perhaps this assumption is not essential for the method.

(4) *Martingale method.* This powerful method has been applied so far only in the case when the c_n are boxes, but it is likely to be valuable in much greater generality. Zhang (1998) has used this approach to prove a central limit theorem for K_n when $p = p_c$ and $d \geq 2$.

Related central limit theorems have been proved by Brånvall (1980) and Grimmett (1981b), but we shall say nothing about these here.

Here are some applications of the central limit theorem. We assume throughout that $\mathbf{c} = (c_n : n \geq 1)$ is a sequence of circuits satisfying $|\Lambda(c_n)| \to \infty$ as $n \to \infty$. For our first example, we define

$$f_n(C(x)) = \begin{cases} 1 & \text{if } |C(x)| = \infty, \\ 0 & \text{if } |C(x)| < \infty, \end{cases}$$

so that $Z_n = \sum_{x \in \Lambda(c_n)} f_n(C(x))$ is the number of vertices in $\Lambda(c_n)$ which lie in infinite open clusters. It is not difficult to verify the conditions of the theorem when $\frac{1}{2} < p < 1$, and we deduce that the number of vertices of $\Lambda(c_n)$ belonging to the infinite open cluster satisfies the central limit theorem.

For our second example we define

$$f_n(C(x)) = \begin{cases} 1 & \text{if } x \leftrightarrow \partial\Lambda(c_n), \\ 0 & \text{otherwise}, \end{cases}$$

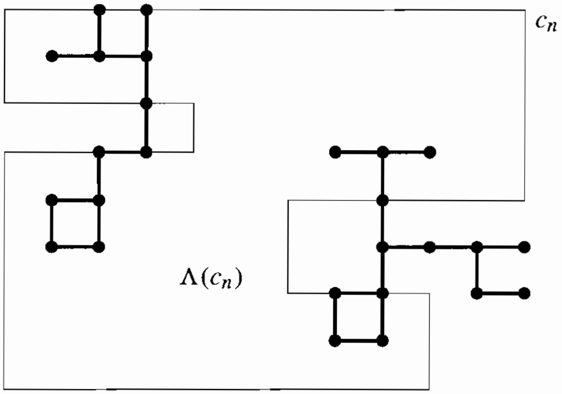


Figure 11.18. All the vertices of the left-hand cluster belong to $\Lambda(c_n)$ and yet its intersection with $\Lambda(c_n)$ is disconnected. The right-hand cluster has a similar property.

so that Z_n is the number of vertices of $\Lambda(c_n)$ which are joined to its boundary by open paths. Once again the conditions of the theorem are easily verified when $p > \frac{1}{2}$, and the asymptotic normality of Z_n follows in this case. If $p < \frac{1}{2}$ then Z_n is a 'boundary effect' rather than a 'volume effect', and so condition (11.66) is not generally valid. Nevertheless, the method of moments may be used to show the asymptotic normality of Z_n in this case also (see Cox and Grimmett (1981)). It may be shown that there exist positive functions $A_1(p)$ and $A_2(p)$ such that

$$A_1(p)|\partial\Lambda(c_n)| \le E_p(Z_n),\ \mathrm{var}_p(Z_n) \le A_2(p)|\partial\Lambda(c_n)| \qquad \text{if } 0 < p < \tfrac{1}{2},$$
$$A_1(p)|\Lambda(c_n)| \le E_p(Z_n),\ \mathrm{var}_p(Z_n) \le A_2(p)|\Lambda(c_n)| \qquad \text{if } \tfrac{1}{2} < p < 1.$$

Cast in the framework of a porous stone immersed in a bucket of water, these observations amount to quantified versions of the statement that the wetting of interior vertices is a boundary effect when $p < p_c$ $(= \frac{1}{2})$ and a volume effect when $p > p_c$.

We turn next to the number of open clusters contained in a large region of the lattice. We need to be careful when counting such clusters, since there may exist open clusters of the lattice whose intersections with $\Lambda(c_n)$ are disconnected. It can even be the case that there exists an open cluster of the lattice all of whose vertices belong to $\Lambda(c_n)$ but whose intersection with $\Lambda(c_n)$ is disconnected; see Figure 11.18 for illustrations of some of these difficulties. As a starting point we define

$$(11.68) \qquad F_n = \sum_{x\in\Lambda(c_n)} \frac{1}{|C(x)\cap\Lambda(c_n)|} I_{\{|C(x)|<\infty\}},$$

and it is not difficult to see that F_n is the number of finite open clusters of the lattice which intersect $\Lambda(c_n)$. The functions in (11.68) are certainly bounded, as well as being constant on infinite graphs. We require a calculation in order to check (11.66) (see Cox and Grimmett (1984, p. 524)), and the asymptotic normality of F_n follows when $p \neq \frac{1}{2}$. If $p < \frac{1}{2}$ then all open clusters are finite (almost surely), so that F_n equals (almost surely) the total number of open clusters of $\mathbb{L}^2$ which intersect $\Lambda(c_n)$. If $p > \frac{1}{2}$, there exists (almost surely) an infinite open cluster, so that the total number of open clusters intersecting $\Lambda(c_n)$ equals either F_n or $F_n + 1$; thus the central limit theorem is valid for this quantity also.

Greater difficulty arises in the study of

$$K_n = \sum_{x \in \Lambda(c_n)} \frac{1}{|C_n(x)|},$$

the number of open clusters of $\Lambda(c_n)$. If $p < \frac{1}{2}$, then Theorem (11.63) is applicable, and we obtain the asymptotic normality of K_n. On the other hand, $f_n(C(x)) = |C_n(x)|^{-1}$ is not constant on infinite graphs, and so the theorem is inapplicable if $p > \frac{1}{2}$. Resorting to a direct application of the method of moments, Cox and Grimmett (1984) have shown the asymptotic normality of K_n when $\frac{1}{2} < p < 1$, so long as the regions $\Lambda(c_n)$ do not have too many bottlenecks. We do not state their theorem here, and neither shall we state precisely the geometry required of the c_n. We note, however, that the conclusion is valid whenever c_n is the boundary of the rectangle $[0, k_1(n)] \times [0, k_2(n)]$ for sequences $(k_1(n) : n \geq 1)$ and $(k_2(n) : n \geq 1)$ satisfying

$$\liminf_{n\to\infty} \left\{ \frac{k_1(n)}{\log(k_1(n)k_2(n))} \right\} = \liminf_{n\to\infty} \left\{ \frac{k_2(n)}{\log(k_1(n)k_2(n))} \right\} = \infty.$$

The martingale method may be used to obtain related results for percolation in general dimensions, subject to certain assumptions of regularity on the regions under study.

11.7 Open Circuits in Annuli

In studying the geometry of the open clusters of percolation on $\mathbb{L}^2$, it is convenient to have simple recipes for constructing certain types of open paths; the probabilities of combinations of such paths may then be estimated with the aid of the FKG and BK inequalities. We have made extensive use already of open crossings of rectangles, and indeed it was by combining such paths in ingenious ways that Russo (1978) and Seymour and Welsh (1978) provided one of the principal foundations of Kesten's proof that $p_c = \frac{1}{2}$ for the square lattice. No longer is the major part of the Russo–Seymour–Welsh technology essential for this result, but such techniques retain their importance for at least two reasons quite apart from historical and intrinsic interest. First, they provide standard tools for approaching

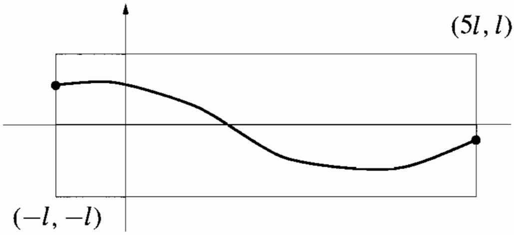


Figure 11.19. A sketch of the event LR(3l, l).

certain two-dimensional processes which are similar to percolation but whose detailed behaviours are not understood so well. Secondly, they enable us to prove certain inequalities such as

(11.69) $$\theta(p) \le A(p - \tfrac{1}{2})^{\alpha} \qquad \text{for } p \ge p_c,$$

where A and α are positive constants. Such inequalities have implications for critical exponents, and we shall discuss this matter in greater detail in the next section.

Of the many results dealing with open crossings of rectangles, the Russo–Seymour–Welsh (RSW) theorem is one of the most fundamental. We begin with some jargon. A *left–right* (respectively *top–bottom*) *crossing* of the rectangle B is an open path in B which joins some vertex on the left (respectively upper) side of B to some vertex on the right (respectively lower) side of B but which uses no edge joining two vertices in the boundary ∂B of B; we require the last condition in order to overcome a mild technical difficulty later. For positive integers k and l, we define the rectangle

$$B(kl, l) = \big[-l, (2k-1)l\big] \times [-l, l];$$

thus $B(l, l) = B(l)$, and $B(kl, l)$ is a rectangle with dimensions $2kl$ by $2l$. We write $\mathrm{LR}(kl, l)$ for the event that there exists a left–right crossing of $B(kl, l)$, and we abbreviate $\mathrm{LR}(l, l)$ to $\mathrm{LR}(l)$; see Figure 11.19. Let $A(l)$ be the annulus $B(3l)\setminus B(l)$, and let $O(l)$ be the event that there exists an open circuit of $A(l)$ containing the origin of the lattice in its interior; see Figure 11.20. A lasting contribution of Russo, Seymour, and Welsh is a proof that, if there is a significant probability of an open crossing of the box $B(l)$, then there is a significant probability of an open circuit in the annulus $A(l)$.

(11.70) Theorem. RSW theorem. *If* $P_p(\mathrm{LR}(l)) = \tau$ *then*

(11.71) $$P_p(O(l)) \ge \Big\{\tau\big(1 - \sqrt{1-\tau}\big)^4\Big\}^{12}.$$

Our principal application is to the case $p = p_c$ $(= \frac{1}{2})$. We have from Lemma (11.21) that there is probability $\frac{1}{2}$ of an open path in the rectangle $[0, n+1] \times [0, n]$

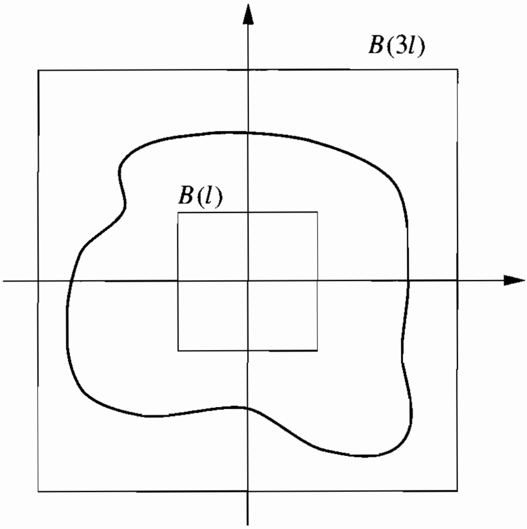


Figure 11.20. An open circuit in the annulus $A(l) = B(3l)\backslash B(l)$, containing the origin in its interior.

joining the left side to the right side, when $p = \frac{1}{2}$. Such a path contains a left–right crossing of the rectangle $[0, n+1] \times [-1, n+1]$. Writing x for the lowest right endvertex of all such crossings, we have that there is probability $\frac{1}{2}$ that the edge leading rightwards out of x is open; therefore, the probability of a left–right crossing of the square $[0, n+2] \times [-1, n+1]$ is at least $\frac{1}{4}$ for $n \geq 0$, and we have proved that $P_{1/2}(\mathrm{LR}(l)) \geq \frac{1}{4}$ for all $l \geq 1$. We deduce from the RSW theorem that

$$(11.72) \qquad P_{1/2}(O(l)) \geq 2^{-24}\left(1 - \tfrac{1}{2}\sqrt{3}\right)^{48} \qquad \text{for all } l \geq 1.$$

The actual number on the right side here is immaterial; the important thing is that $P_{1/2}(O(l))$ is bounded away from zero uniformly in l. We shall see applications of (11.72) in the next section.

The RSW theorem is proved in several stages, of which the following is the hardest.

(11.73) Lemma. *If $P_p(\mathrm{LR}(l)) = \tau$ then*

$$(11.74) \qquad P_p\left(\mathrm{LR}(\tfrac{3}{2}l, l)\right) \geq \left(1 - \sqrt{1-\tau}\right)^3.$$

The remaining work is contained in the next lemma.

(11.75) Lemma. *It is the case that*

$$P_p\big(\mathrm{LR}(2l,l)\big) \geq P_p\big(\mathrm{LR}(l)\big)P_p\big(\mathrm{LR}(\tfrac{3}{2}l,l)\big)^2, \tag{11.76}$$

$$P_p\big(\mathrm{LR}(3l,l)\big) \geq P_p\big(\mathrm{LR}(l)\big)P_p\big(\mathrm{LR}(2l,l)\big)^2, \tag{11.77}$$

$$P_p(O(l)) \geq P_p\big(\mathrm{LR}(3l,l)\big)^4. \tag{11.78}$$

The RSW theorem is an immediate consequence of these two lemmas. We turn next to its proof, and we begin with Lemma (11.73). We shall make repeated use of the square root trick; recall from (11.14) that, if A_1 and A_2 are increasing events having equal probability, then

$$P_p(A_1) \geq 1 - \sqrt{1 - P_p(A_1 \cup A_2)}. \tag{11.79}$$

Proof of Lemma (11.73). We follow Russo (1981). Suppose that $0 \leq p \leq 1$ and consider percolation on $\mathbb{L}^2$ with edge-probability p. The target is to build a left–right crossing of the rectangle $B(\frac{3}{2}l, l)$ using open crossings of $2l \times 2l$ squares. In doing this, we shall encounter certain problems arising from the dependence between certain events; in order to deal with these problems, we shall make use of the idea of the 'lowest' open crossing of a rectangle. Let $\mathcal{C}$ be the collection of all left–right crossings of $B(l)$, and let Π be the lowest such crossing. Some topology is required in order to show that Π exists and is defined uniquely whenever $\mathcal{C}$ is non-empty, but the reader may be prepared to accept the following argument. There is a natural partial order on the set of left–right crossings of $B(l)$: for two such crossings π_1 and π_2 we write $\pi_1 \leq \pi_2$ if π_1 is contained in the closed bounded region of $B(l)$ lying beneath π_2. By geometrical considerations (see Figure 11.21), the union of the edge sets of any two crossings π_1 and π_2 contains a crossing π_3 such that $\pi_3 \leq \pi_1$ and $\pi_3 \leq \pi_2$. It follows that the union of the edge sets of all open crossings of $B(l)$ contains an open crossing Π which satisfies $\Pi \leq \pi$ for all open crossings π, and Π is called the *lowest* left–right crossing. We shall use these ideas in the following way: first, there exists a *unique* lowest left–right crossing whenever there exists a left–right crossing, and secondly, the value taken by the lowest left–right crossing Π is independent of the states of edges of $B(l)$ having at least one endvertex in the subset of $B(l)$ lying strictly above Π. This latter property of Π is crucial and requires some elucidation. Let h be a (possibly non-open) path which traverses $B(l)$ from left to right, and consider the event $\{\Pi = h\}$ that h is the lowest left–right crossing. This event depends only on the states of edges in h and below h, since $\Pi = h$ if and only if h is open and there exists no other left–right crossing Π' such that $\Pi' \leq \Pi$. Thus Π enjoys a property similar to that of a stopping time for a Markov process.

Let $\mathcal{T}$ be the set of paths of $B(l)$ which traverse $B(l)$ from left to right but which use no edge contained in the boundary of $B(l)$. Each path $\pi \in \mathcal{T}$ cuts $[-l, l] \times [-l, l]$ into two parts, being the simply connected subsets of $\mathbb{R}^2$ lying above and beneath π. For $\pi \in \mathcal{T}$, we denote by $(0, y_\pi)$ the last vertex encountered

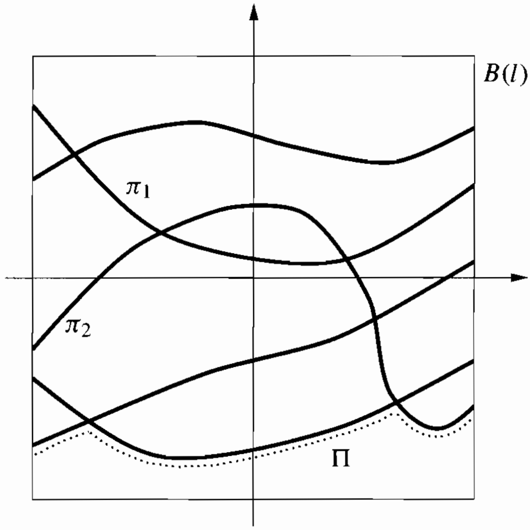


Figure 11.21. Note that the union of any two left–right crossings π_1 and π_2 contains a crossing π_3 satisfying $\pi_3 \leq \pi_1$ and $\pi_3 \leq \pi_2$. The dotted crossing is the lowest left–right crossing of the box.

on the vertical axis of $\mathbb{L}^2$ when π is traversed from left to right, and we write $\mathcal{T}^-$ (respectively $\mathcal{T}^+$) for the subset of $\mathcal{T}$ containing paths π for which $y_\pi \leq 0$ (respectively $y_\pi \geq 0$). For $\pi \in \mathcal{T}$, we denote by π_r the sub-path of π joining $(0, y_\pi)$ to the right side of $B(l)$. Let π_r' be the reflection of π_r in the line $\{(l, k) : -\infty < k < \infty\}$, and let $\pi_\mathrm{r} \cup \pi_\mathrm{r}'$ be the path obtained by prolonging π_r with π_r'; the path $\pi_\mathrm{r} \cup \pi_\mathrm{r}'$ traverses the square $B(l)' = [0, 2l] \times [-l, l]$ from left to right. See Figure 11.22 for an illustration of this notation. For any $\pi \in \mathcal{T}$, we write $U(\pi)$ for the set of vertices of $B(l)'$ lying in or above the path $\pi_\mathrm{r} \cup \pi_\mathrm{r}'$, together with all edges of $B(l)'$ joining pairs of such vertices, at least one of which lies strictly above $\pi_\mathrm{r} \cup \pi_\mathrm{r}'$.

We are interested in the following events. Let L^- (respectively L^+) be the event that there exists an open path π in $\mathcal{T}^-$ (respectively $\mathcal{T}^+$), which is to say that $L^- = \{y_\Pi \leq 0\}$. For $\pi \in \mathcal{T}$, we denote by L_π the event that π is the lowest left–right crossing Π of $B(l)$, and by A_π the event that π is open. For $\pi \in \mathcal{T}$, we write M_π for the event that there exists an open path in $U(\pi)$ joining some vertex on the upper side of $B(l)'$ to some vertex of $\pi_\mathrm{r} \cup \pi_\mathrm{r}'$, and we write M_π^- (respectively M_π^+) for the event that such a path exists having its second endvertex in π_r (respectively π_r'). Finally, we write N for the event that there exists a left–right crossing of $B(l)'$, and N^- (respectively N^+) for the event that such a path exists having its first endvertex on the non-positive (respectively non-negative) part of the vertical axis.

We shall use these events in the following way. Suppose that, for some $\pi \in \mathcal{T}^-$, the event $A_\pi \cap M_\pi^- \cap N^+$ occurs, and consult Figure 11.22 for a sketch of this

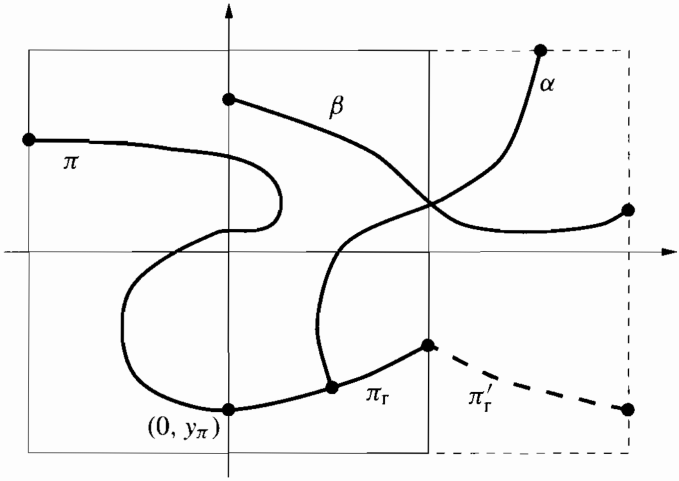


Figure 11.22. The path π crosses $B(l)$ from left to right. The path α guarantees the occurrence of the event M_π^-, and the path β guarantees the occurrence of N^+.

event. It is clear from the figure that this event necessarily entails the existence of a left–right crossing of $B(\frac{3}{2}l, l)$, so that

$$P_p\big(\mathrm{LR}(\tfrac{3}{2}l, l)\big) \geq P_p\left(N^+ \cap \left\{\bigcup_{\pi \in \mathcal{T}^-} [A_\pi \cap M_\pi^-]\right\}\right); \tag{11.80}$$

thus we seek a lower bound for the right side of this inequality. We note first that N^+ is an increasing event. Furthermore,

$$G = \bigcup_{\pi \in \mathcal{T}_-} [A_\pi \cap M_\pi^-]$$

is the event that there exists an open path π in $\mathcal{T}^-$ for which M_π^- occurs; this is an increasing event also, giving by the FKG inequality that

$$P_p\big(\mathrm{LR}(\tfrac{3}{2}l, l)\big) \geq P_p(N^+)P_p(G). \tag{11.81}$$

We apply the square root trick (11.79) to the events N^+ and N^- to find that

$$\begin{aligned} P_p(N^+) &\geq 1 - \sqrt{1 - P_p(N^+ \cup N^-)} \\ &= 1 - \sqrt{1 - \tau}, \end{aligned} \tag{11.82}$$

where $\tau = P_p(\mathrm{LR}(l))$. Now,

$$G \supseteq \bigcup_{\pi \in \mathcal{T}^-} [L_\pi \cap M_\pi^-],$$

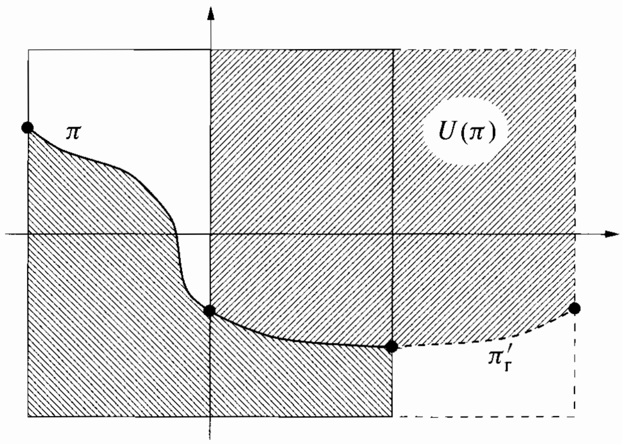


Figure 11.23. The two hatched regions indicate the disjoint sets of edges whose states determine the occurrence or not of M_π^- and L_π.

a union of disjoint events, and therefore

(11.83) $$P_p(G) \geq \sum_{\pi \in \mathcal{T}^-} P_p(M_\pi^- \mid L_\pi) P_p(L_\pi).$$

We claim that

(11.84) $$P_p(M_\pi^- \mid L_\pi) \geq 1 - \sqrt{1-\tau} \qquad \text{for all } \pi \in \mathcal{T};$$

once this claim has been proved, the result will follow fairly quickly as follows. We substitute (11.82)–(11.84) into (11.80) to obtain

(11.85) $$P_p\big(\mathrm{LR}(\tfrac{3}{2}l, l)\big) \geq \big\{1 - \sqrt{1-\tau}\big\}^2 \sum_{\pi \in \mathcal{T}^-} P_p(L_\pi).$$

However,

$$\sum_{\pi \in \mathcal{T}_-} P_p(L_\pi) = P_p(L^-);$$

furthermore, we may apply the square root trick (11.79) to the events L^- and L^+ to obtain

$$\begin{aligned} P_p(L^-) &\geq 1 - \sqrt{1 - P_p(L^- \cup L^+)} \\ &= 1 - \sqrt{1-\tau}, \end{aligned}$$

which, when substituted into (11.85), yields the conclusion of the lemma.

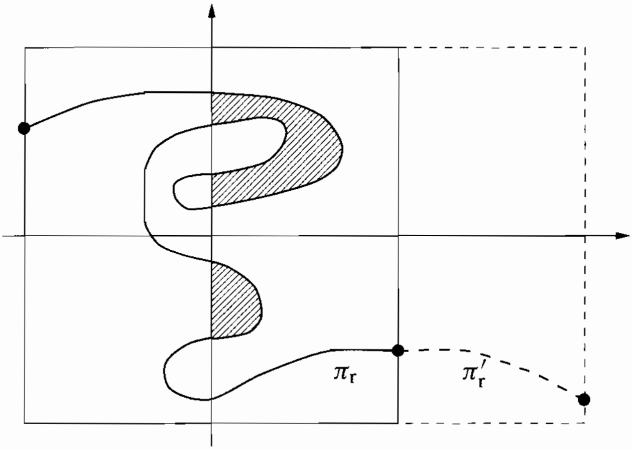


Figure 11.24. The set J is the set of edges in the hatched region.

It remains to prove (11.84), for which we shall use the square root trick once again. Suppose that $\pi \in \mathcal{T}$. As an illustration of the argument, we consider first the case when π intersects the vertical axis of $\mathbb{L}^2$ once only. In this case, the events M_π^- and L_π are independent, since M_π^- is defined in terms of the states of edges of $B(l)'$ with at least one endvertex strictly above $\pi_r \cup \pi'_r$, whereas L_π is defined in terms of the states of edges in and below π in $B(l)$; see Figure 11.23. Thus

(11.86) $$P_p(M_\pi^- \mid L_\pi) = P_p(M_\pi^-)$$

in this case; we now apply the square root trick to the events M_π^- and M_π^+ to deduce that

(11.87) $$\begin{aligned} P_p(M_\pi^-) &\geq 1 - \sqrt{1 - P_p(M_\pi^- \cup M_\pi^+)} \\ &= 1 - \sqrt{1 - P_p(M_\pi)} \\ &\geq 1 - \sqrt{1 - \tau} \end{aligned}$$

as required, since M_π occurs whenever $B(l)'$ contains a top–bottom crossing. We now lift the restriction that π intersect the vertical axis once only. Let H be the set of edges of π which lie in $U(\pi)$, and let J be the set of edges of $U(\pi)$ which have at least one endvertex in the simply connected subset of $[-l, l] \times [-l, l]$ ($\subseteq \mathbb{R}^2$) lying strictly beneath π (see Figure 11.24).

We claim that

$$P_p(M_\pi^- \mid L_\pi) = P_p(M_\pi^- \mid J \cup H \text{ open}),$$

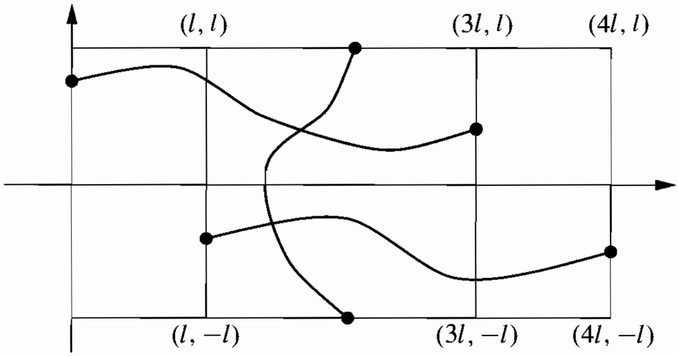


Figure 11.25. Three open paths which guarantee the occurrence of the events LR_1, LR_2, and TB_1. Their union contains a left–right crossing of $[0, 4l] \times [-l, l]$.

where $\{J \cup H \text{ open}\}$ is the event that every edge in $J \cup H$ is open. To see this, note from Figure 11.24 that, if π is an open crossing of $B(l)$, we do not have to know the states of edges in J in determining whether or not M_π^- occurs. Now, M_π^- and $\{J \cup H \text{ open}\}$ are increasing events, and we apply the FKG inequality to find that

$$P_p(M_\pi^- \mid L_\pi) \geq P_p(M_\pi^-) \qquad \text{for all } \pi \in \mathcal{T}.$$

Inequality (11.87) is valid for all $\pi \in \mathcal{T}$, and we obtain inequality (11.84) as before. □

Proof of Lemma (11.75). We follow Russo (1981) again, beginning with (11.76). Let LR_1 (respectively LR_2) be the event that there exists a left–right crossing of the rectangle $[0, 3l] \times [-l, l]$ (respectively $[l, 4l] \times [-l, l]$) and let TB_1 be the event that there exists a top–bottom crossing of the square $[l, 3l] \times [-l, l]$; see Figure 11.25. If all three such paths exist, there exists a left–right crossing of $[0, 4l] \times [-l, l]$, so that

$$\begin{aligned} P_p\big(\mathrm{LR}(2l, l)\big) &\geq P_p(\mathrm{LR}_1 \cap \mathrm{LR}_2 \cap \mathrm{TB}_1) \\ &\geq P_p(\mathrm{LR}_1) P_p(\mathrm{LR}_2) P_p(\mathrm{TB}_1) \end{aligned}$$

by the FKG inequality, and we have proved (11.76).

Inequality (11.77) follows similarly, by considering the events

$$\begin{aligned} \mathrm{LR}_3 &= \big\{\text{there exists a left–right crossing of } [0, 4l] \times [-l, l]\big\}, \\ \mathrm{LR}_4 &= \big\{\text{there exists a left–right crossing of } [2l, 6l] \times [-l, l]\big\}, \\ \mathrm{TB}_2 &= \big\{\text{there exists a top–bottom crossing of } [2l, 4l] \times [-l, l]\big\}; \end{aligned}$$

see Figure 11.26.

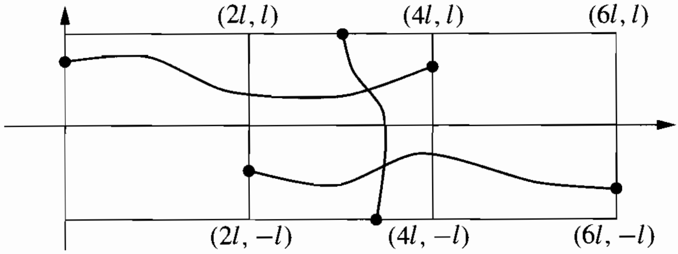


Figure 11.26. Three open paths which guarantee the occurrence of LR_3, LR_4, and TB_2. Their union contains a left–right crossing of $[0, 6l] \times [-l, l]$.

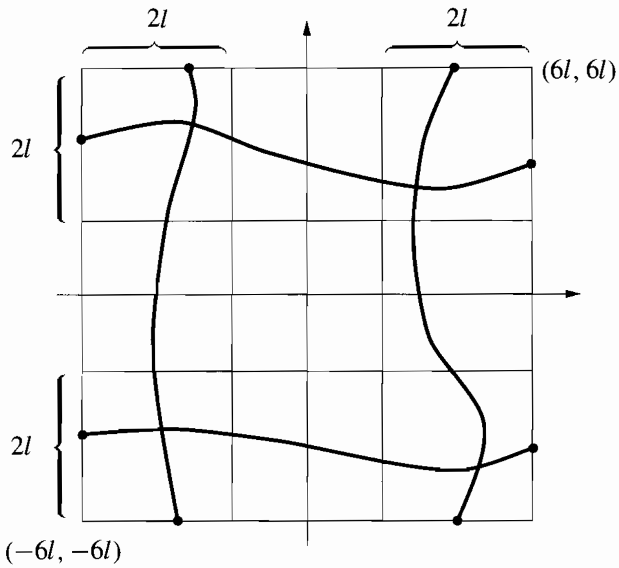


Figure 11.27. If the four $2l \times 6l$ rectangles are traversed by open paths as indicated, then there exists an open circuit in the annulus $A(l)$ containing the origin in its interior.

It is equally easy to prove inequality (11.78). We express $A(l)$ as the union of the four rectangles having dimensions $2l \times 6l$ illustrated in Figure 11.27. Each of these rectangles is traversed by an open path joining its two shorter sides with probability at least $P_p(\mathrm{LR}(3l, l))$, and the result follows by the FKG inequality. Note that we have used the convention that 'crossings' of rectangles use no edges in the boundary. □

11.8 Power Law Inequalities

The greatest current mystery in percolation is the behaviour of such quantities as θ and χ at and near the critical point. We recall such conjectures as

$$\theta(p) \approx (p - p_c)^{\beta} \qquad \text{as } p \downarrow p_c,$$
$$\chi(p) \approx (p_c - p)^{-\gamma} \qquad \text{as } p \uparrow p_c,$$
$$P_{p_c}(|C| = n) \approx n^{-1-1/\delta} \qquad \text{as } n \to \infty,$$

for percolation on $\mathbb{L}^d$ when $d \geq 2$. No rigorous proofs of such relations are known at present. On the other hand, the methods of Chapters 9 and 10 may be used to obtain certain inequalities of the right form: for example, we saw in Section 10.2 that $\delta \geq 2$ whenever δ exists, which fact suggests the possible validity of an inequality of the form

(11.88) $$P_{p_c}(|C| \geq n) \geq An^{-1/2}$$

for some positive constant $A = A(d)$. In the absence of rigorous proof of power law behaviour, we may seek power law *inequalities* of the form of (11.88). As usual, the case when $d = 2$ is the most tractable, since it is in this case that the technique of duality is most easily available. We describe such results in this section.

We consider percolation on $\mathbb{L}^2$, for which the critical probability is $p_c = \frac{1}{2}$, and $\theta(\frac{1}{2}) = 0$. The following theorem contains our principal estimates at the critical point.

(11.89) Theorem. Power law inequalities at the critical point. *There exist positive finite constants A_i, α_i, for $i \geq 1$, such that the following statements are valid*:

(11.90) $$\tfrac{1}{2}n^{-1/2} \leq P_{1/2}\big(0 \leftrightarrow \partial B(n)\big) \leq A_1 n^{-\alpha_1} \qquad \textit{for all } n \geq 1,$$

(11.91) $$\tfrac{1}{2}n^{-1/2} \leq P_{1/2}(|C| \geq n) \leq A_2 n^{-\alpha_2} \qquad \textit{for all } n \geq 1,$$

(11.92) $$E_{1/2}(|C|^{\alpha_3}) < \infty.$$

We have as consequences of (11.90) and (11.91) that the critical exponents ρ and δ satisfy $2 \leq \rho < \infty$ and $2 \leq \delta < \infty$ in two dimensions, if ρ and δ exist. Better inequalities than these are available: see Kesten (1987b) for proofs that $\rho \geq 3$ and $\delta \geq 5$. The inequality $\delta \geq 2$ is valid in all dimensions (see Proposition (10.29)), but we present here an alternative proof for two dimensions.

The principal components of Theorem (11.89) are the inequalities (11.90) concerning the probability of an open path from the origin to the surface of the box $B(n)$; the left-hand inequality is proved by the BK inequality, and the right-hand inequality by an application of the RSW theorem. We shall prove Theorem (11.89) later in this section. We turn next to power law inequalities for θ, χ, and χ^{f} when p is near to its critical value $\frac{1}{2}$.

(11.93) Theorem. Power law inequalities for θ, χ, and χ^{f}. *There exist positive finite constants A_i, α_i, for $i \geq 1$, such that the following statements are valid:*

$$\theta(p) \leq A_4(p - \tfrac{1}{2})^{\alpha_4} \qquad \text{if } p > \tfrac{1}{2}, \tag{11.94}$$

$$\chi(p) \leq A_5(\tfrac{1}{2} - p)^{-\alpha_5} \qquad \text{if } p < \tfrac{1}{2}, \tag{11.95}$$

$$A_6(p - \tfrac{1}{2})^{-1/4} \leq \chi^{\mathrm{f}}(p) \leq A_7(p - \tfrac{1}{2})^{-\alpha_7} \qquad \text{if } p > \tfrac{1}{2}. \tag{11.96}$$

There are corresponding lower bounds for θ and χ which are valid in all dimensions: we have from Theorem (5.8) and equation (10.41) that there exists a (> 0) such that

$$\theta(p) \geq a(p - \tfrac{1}{2}) \qquad \text{if } p > \tfrac{1}{2} \tag{11.97}$$

and

$$\chi(p) \geq \tfrac{1}{4}(\tfrac{1}{2} - p)^{-1} \qquad \text{if } p < \tfrac{1}{2}.$$

Likewise, the lower bound in (11.96) for χ^{f} may be improved using arguments of Newman (1986, 1987b, c), subject possibly to a minor assumption concerning the asymptotic behaviour of the cluster size distribution. It is possible to use the arguments in the proofs of the above two theorems to show the existence of positive constants A_8 and α_8 such that

$$P_p\big(n \leq |C| < \infty\big) \leq A_8 n^{-\alpha_8} \qquad \text{for all } n \geq 1 \text{ and } 0 \leq p \leq 1; \tag{11.98}$$

see Kesten (1981). This inequality has an immediate application to the differentiability of the number $\kappa(p)$ of open clusters per vertex. We know from Theorem (6.108) that κ is analytic on $[0, \frac{1}{2})$, and it follows from Theorem (11.4) that κ is analytic on $(\frac{1}{2}, 1]$ also. The only lingering challenge is to understand its behaviour when $p = p_{\mathrm{c}} = \frac{1}{2}$. Certainly κ is continuously differentiable at $\frac{1}{2}$, from Theorem (4.31). Expressing $\kappa(p)$ as an infinite series (as in (4.18)) and differentiating twice, term by term, we may use (11.98) to show that the resulting series is uniformly convergent on a neighbourhood of $\frac{1}{2}$; thus κ is twice continuously differentiable at $\frac{1}{2}$. The argument fails to show thrice differentiability, and it is conjectured that the third derivative of κ does not exist at the critical probability (see also the discussion around equation (9.6)).

Proof of Theorem (11.89). We begin with (11.90), of which (11.91) and (11.92) will be easy corollaries. For the left-hand inequality of (11.90) we use the BK inequality, and we follow van den Berg and Kesten (1985). Let $\mathrm{LR}(n)$ be the event that there exists an open path in the rectangle $R_n = [0, 2n] \times [0, 2n - 1]$ joining some vertex on its left side to some vertex on its right side, and recall from Lemma (11.21) that

$$P_{1/2}\big(\mathrm{LR}(n)\big) = \tfrac{1}{2} \qquad \text{for } n \geq 1. \tag{11.99}$$

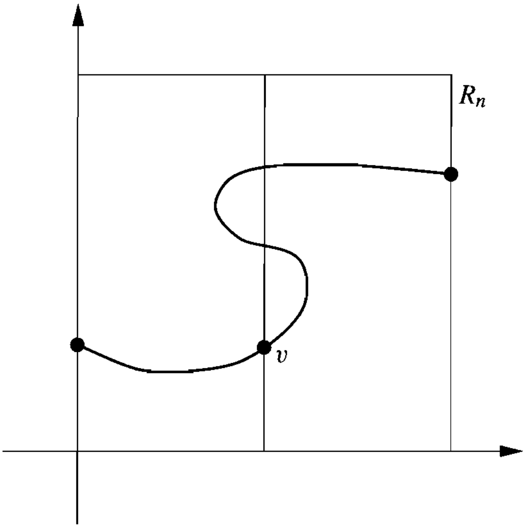


Figure 11.28. Any open path traversing R_n contains some vertex v on the centre line, and this vertex is joined by edge-disjoint open paths to the left and right sides of R_n.

Any such open crossing of R_n intersects its centre line $\{(n,k) : 0 \le k \le 2n-1\}$ in at least one vertex, and any such vertex is connected by two disjoint paths to the left side and to the right side of R_n; see Figure 11.28. Hence

$$P_{1/2}\big(\mathrm{LR}(n)\big) \le \sum_{k=0}^{2n-1} P_{1/2}\big(A_n(k) \circ A_n(k)\big),$$

where $A_n(k)$ is the event that the vertex (n,k) is joined by an open path to the surface of the box $B(n, (n,k))$ having side-length $2n$ and centre at (n,k).

We use the BK inequality to find that

$$\begin{aligned} P_{1/2}\big(\mathrm{LR}(n)\big) &\le \sum_{k=0}^{2n-1} P_{1/2}\big(A_n(k)\big)^2 \\ &= 2n P_{1/2}\big(0 \leftrightarrow \partial B(n)\big)^2, \end{aligned}$$

and hence

$$P_{1/2}\big(0 \leftrightarrow \partial B(n)\big) \ge \tfrac{1}{2} n^{-1/2}$$

by (11.99), as required.

For the upper bound on $P_{1/2}(0 \leftrightarrow \partial B(n))$ we follow Kesten (1981), using the RSW theorem and particularly (11.72): there exists $\zeta > 0$ such that

$$(11.100) \qquad P_{1/2}(O(l)) \ge \zeta \qquad \text{for all } l \ge 1,$$

where $O(l)$ is the event that the annulus $A(l) = B(3l)\backslash B(l)$ contains an open circuit having the origin in its interior. We use the dual lattice technique here, writing $\mathbb{L}_d^2$ for the dual of $\mathbb{L}^2$ as before. Let $B(k)_d$ be the square $B(k) + (\frac{1}{2}, \frac{1}{2})$ in the dual lattice, and let $A(l)_d = B(3l)_d \setminus B(l)_d$ be the corresponding annulus in the dual. Let $O(l)_d$ be the event that $A(l)_d$ contains a closed circuit of the dual lattice having the dual origin $0_d = (\frac{1}{2}, \frac{1}{2})$ in its interior. We have from (11.100) that

$$P_{1/2}(O(l)_d) \geq \zeta \qquad \text{for all } l \geq 1, \tag{11.101}$$

since each dual edge is closed with probability $\frac{1}{2}$. Suppose that there exists an open path from the origin of $\mathbb{L}^2$ to the surface of the box $B(3^k+1)$. Such a path traverses each of the dual annuli $A(3^r)_d$, for $0 \leq r < k$, and therefore none of the events $O(3^r)_d$ occurs for $0 \leq r < k$. However, the events $\{O(3^r)_d : 0 \leq r < k\}$ are defined in terms of disjoint sets of dual edges, and they are therefore independent. Therefore,

$$\begin{aligned} P_{1/2}\big(0 \leftrightarrow \partial B(3^k+1)\big) &\leq P_{1/2}\big(O(3^r)_d \text{ does not occur for } 0 \leq r < k\big) \\ &\leq (1-\zeta)^k \end{aligned}$$

by (11.101). If $3^k + 1 \leq n < 3^{k+1} + 1$ then

$$\begin{aligned} P_{1/2}\big(0 \leftrightarrow \partial B(n)\big) &\leq P_{1/2}\big(0 \leftrightarrow \partial B(3^k+1)\big) \\ &\leq (1-\zeta)^k \\ &\leq A_1 n^{-\alpha_1} \end{aligned} \tag{11.102}$$

for some positive finite constants A_1 and α_1 depending on ζ only. This completes the proof of (11.90).

It is not difficult to deduce (11.91) from (11.90). Clearly

$$P_{1/2}\big(0 \leftrightarrow \partial B(n)\big) \leq P_{1/2}(|C| \geq n) \leq P_{1/2}\big(0 \leftrightarrow \partial B(\lfloor \tfrac{1}{2}\sqrt{n} - 1 \rfloor)\big)$$

since $|C| \geq n$ if $0 \leftrightarrow \partial B(n)$, and a cluster of size n cannot be accommodated within $B(m)$ if $n > |B(m)| = (2m+1)^2$; (11.91) follows immediately. Similarly, if $0 < a < 1$ then

$$\begin{aligned} E_{1/2}(|C|^a) &= \sum_{n=1}^{\infty} \{n^a - (n-1)^a\} P_{1/2}(|C| \geq n) \\ &\leq A \sum_{n=1}^{\infty} n^{a-1} n^{-\alpha_2} \end{aligned}$$

for some finite constant A, by (11.91). This summation is finite if $a < \alpha_2$. □

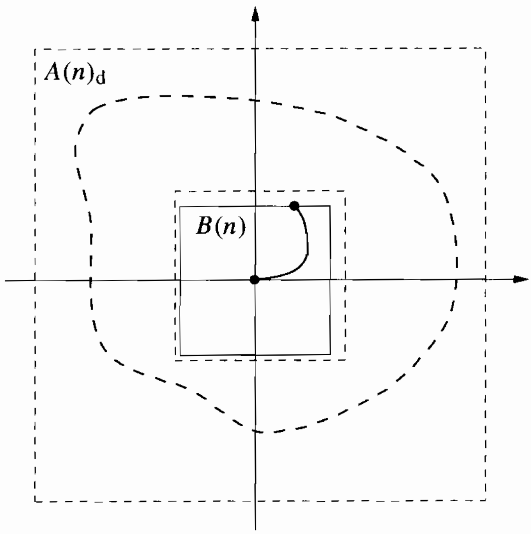


Figure 11.29. If there exists an open path from the origin to $\partial B(n)$ together with a closed circuit of $A(n)_d$ surrounding the origin, then the open cluster at the origin is finite with size no less than n.

Proof of Theorem (11.93). We follow largely Kesten (1981, 1982), and begin with the upper bound for θ. We observe first that

(11.103) $$\theta(p) \le P_p\big(0 \leftrightarrow \partial B(n)\big) \qquad \text{for all } n,$$

since any infinite open path from the origin contains vertices of $\partial B(n)$. We use inequality (2.31) to relate $P_p(0 \leftrightarrow \partial B(n))$ to $P_{1/2}(0 \leftrightarrow \partial B(n))$. The number of edges in $B(n)$ is no greater than $2(2n+1)^2 \le 18n^2$, and therefore

(11.104)
$$P_{p_1}\big(0 \leftrightarrow \partial B(n)\big) \ge (p_1/p_2)^{18n^2} P_{p_2}\big(0 \leftrightarrow \partial B(n)\big) \qquad \text{if } 0 \le p_1 < p_2 \le 1.$$

Setting $p_1 = \frac{1}{2}$ and $p_2 = p > \frac{1}{2}$, we obtain

(11.105) $$P_{1/2}\big(0 \leftrightarrow \partial B(n)\big) \ge (2p)^{-18n^2} P_p\big(0 \leftrightarrow \partial B(n)\big).$$

When combined with (11.103) and (11.90) this implies that, for $p > \frac{1}{2}$,

$$\theta(p) \le A_1(2p)^{18n^2} n^{-\alpha_1} \qquad \text{for all } n \ge 1.$$

We now choose n by

(11.106) $$n = \big\lfloor \{\log(2p)\}^{-1/2} \big\rfloor,$$

noting that $n = n(p) \sim \{2(p - \frac{1}{2})\}^{-1/2}$ as $p \downarrow \frac{1}{2}$, to obtain

(11.107) $$\begin{aligned} \theta(p) &\le A_1 e^{18} n^{-\alpha_1} \\ &\sim A_4(p - \tfrac{1}{2})^{\alpha_1/2} \qquad \text{as } p \downarrow \tfrac{1}{2}, \end{aligned}$$

for some constant A_4, and the result follows.

We turn next to the lower bound for χ^{f}, and we suppose that $p > \frac{1}{2}$. It is the case that $n \le |C| < \infty$ if there exists both an open path from the origin to the surface of $B(n)$ and a closed circuit of the dual annulus $A(n)_{\mathrm{d}}$ containing the origin of $\mathbb{L}^2$ in its interior; see Figure 11.29. These two events are defined in terms of disjoint sets of edges, and they are therefore independent. Hence

$$\begin{aligned}(11.108)\qquad \chi^{\mathrm{f}}(p) &\ge nP_p\big(n \le |C| < \infty\big)\\ &\ge nP_p\big(0 \leftrightarrow \partial B(n),\, O(n)_{\mathrm{d}}\big)\\ &= nP_p\big(0 \leftrightarrow \partial B(n)\big)P_p(O(n)_{\mathrm{d}})\\ &\ge \tfrac{1}{2}n^{1/2}P_p(O(n)_{\mathrm{d}})\end{aligned}$$

by (11.90), since $P_p\big(0 \leftrightarrow \partial B(n)\big)$ is a non-decreasing function of p. Also,

$$(11.109)\qquad P_p(O(n)_{\mathrm{d}}) = P_{1-p}(O(n))$$

since each edge of the dual lattice is closed with probability $1 - p$. We use inequality (2.31) again to find that

$$P_{1-p}(O(n)) \ge \big\{2(1-p)\big\}^{98n^2} P_{1/2}(O(n)),$$

where $98n^2$ is an upper bound for the number of edges in the annulus $A(n)$. However, $P_{1/2}(O(n)) > \zeta$ (> 0) by (11.100), giving by (11.108) and (11.109) that

$$\chi^{\mathrm{f}}(p) \ge \tfrac{1}{2}\zeta\big\{2(1-p)\big\}^{98n^2} n^{1/2}.$$

We choose n by

$$n = \Big\lfloor \big\{-\log[2(1-p)]\big\}^{-1/2} \Big\rfloor,$$

noting that $n = n(p) \sim \{2(p - \frac{1}{2})\}^{-1/2}$ as $p \downarrow \frac{1}{2}$, to obtain

$$\begin{aligned}\chi^{\mathrm{f}}(p) &\ge \tfrac{1}{2}\zeta e^{-98} n^{1/2}\\ &\sim \tfrac{1}{2}\zeta e^{-98}\big\{2(p - \tfrac{1}{2})\big\}^{-1/4} \qquad \text{as } p \downarrow \tfrac{1}{2}.\end{aligned}$$

It follows that the left-hand inequality in (11.96) is valid for all $p > \frac{1}{2}$ and some constant A_6.

In demonstrating the upper bounds for χ and χ^{f} we deviate somewhat from the original argument of Kesten (1981), making use instead of material in Section 5.2. Let $S(n)$ be the ball $\{x \in \mathbb{Z}^d : \delta(0,x) \le n\}$ with radius n and centre at the origin, and let

$$g_p(k) = P_p\big(0 \leftrightarrow \partial S(k)\big),$$

as in Section 5.2. We have from (5.22) that

$$(11.110)\qquad g_\alpha(n) \le g_\beta(n)\exp\left(-(\beta - \alpha)\left[\frac{n}{\sum_{i=0}^{n} g_\beta(i)} - 1\right]\right) \qquad \text{if } 0 \le \alpha < \beta \le 1.$$

We set $\alpha = p\ (< \frac{1}{2})$ and $\beta = \frac{1}{2}$ to obtain

$$(11.111)\qquad g_p(n) \le \exp\left(-(\tfrac{1}{2} - p)\left[\frac{n}{\sum_{i=0}^{n} g_{1/2}(i)} - 1\right]\right) \quad \text{for } 0 \le p < \tfrac{1}{2}.$$

However,

$$g_{1/2}(i) \le P_{1/2}\big(0 \leftrightarrow \partial B(\lfloor \tfrac{1}{2} i \rfloor)\big)$$

since $B(\lfloor \frac{1}{2} i \rfloor)$ is a subset of $S(i)$. Now, we have from (11.90) that $g_{1/2}(i) \le A_1 i^{-\alpha_1}$ for $i \ge 1$, where $0 < \alpha_1 \le \frac{1}{2}$. We insert this into (11.111) to obtain

$$(11.112)\qquad g_p(n) \le A_9 \exp\big\{-A_{10} n^{\alpha_1} (\tfrac{1}{2} - p)\big\} \qquad \text{if } 0 \le p < \tfrac{1}{2},$$

for all n and for appropriate positive constants A_9, A_{10}. It is a simple matter to extract the required bounds for χ and χ^{f} from this inequality. First we recall the obvious inequality (see (5.7))

$$(11.113)\qquad P_p(|C| > n) \le g_p\big(\lfloor \sqrt{n/\nu} \rfloor\big),$$

where $0 < \nu < \infty$. Hence, for $0 \le p < \frac{1}{2}$,

$$\begin{aligned}\chi(p) &= \sum_{n=0}^{\infty} P_p(|C| > n) \\ &\le \sum_{n=0}^{\infty} A_9 \exp\big\{-A_{11} n^{\alpha_1/2} (\tfrac{1}{2} - p)\big\}\end{aligned}$$

for some positive constant A_{11}. We replace the summation by an integral and make the change of variables $y = n^{\alpha_1/2}(\frac{1}{2} - p)$ to obtain

$$\chi(p) \le A_{12} (\tfrac{1}{2} - p)^{-2/\alpha_1} \int_0^\infty y^{-1+(2/\alpha_1)} e^{-A_{11} y}\, dy$$

for some positive constant A_{12}. The integral converges, and we have shown (11.95).

Finally we prove the upper bound for χ^{f}, and we assume therefore that $\frac{1}{2} < p < 1$. We recall from the discussion prior to (11.3) that there exists a positive constant λ with the following property: if $|C| = n$, the origin lies in the interior of some closed circuit of the dual lattice having length at least $\lambda\sqrt{n}$, and there exists such a circuit containing at least one of the dual vertices $(k + \frac{1}{2}, \frac{1}{2})$ for $0 \le k < n$. Such a dual vertex belongs to a closed cluster of the dual having size at least $\lambda\sqrt{n}$, and it follows that

$$\begin{aligned}P_p(|C| = n) &\le \sum_{k=0}^{n-1} P_{1-p}\big(|C| \ge \lambda\sqrt{n}\big) \\ &\le n g_{1-p}\big(\lfloor A_{13} n^{1/4} \rfloor\big)\end{aligned}$$

by (11.113), for some positive constant A_{13}. Now $1 - p < \frac{1}{2}$, and so

$$g_{1-p}(k) \le A_9 \exp\left\{-A_{10}k^{\alpha_1}(p - \tfrac{1}{2})\right\}$$

by (11.112). We now argue as before to deduce that there exists a positive constant A_{14} such that

$$\begin{aligned}\chi^{\mathrm{f}}(p) &= \sum_{n=1}^{\infty} n P_p(|C| = n) \\ &\le A_{14}(p - \tfrac{1}{2})^{-12/\alpha_1}\end{aligned}$$

as required. □

11.9 Inhomogeneous Square and Triangular Lattices

Most of the methods and results in this book are valid in greater generality than simply bond percolation on the cubic lattice $\mathbb{L}^d$. New difficulties can however arise for percolation models having few or no symmetry properties. For example, whereas the arguments of Chapter 5 use little more than translation invariance, those of Chapter 7 use in addition a certain amount of rotation invariance. We illustrate some of the difficulties which may arise by treating inhomogeneous percolation on the square and triangular lattices. We shall see that the first case is the easier, by virtue of the fact that this model possesses axes of symmetry.

Consider the square lattice $\mathbb{L}^2$, and let $\mathbf{p} = (p_{\mathrm{h}}, p_{\mathrm{v}}) \in [0, 1]^2$. We declare each horizontal (respectively vertical) edge to be *open* with probability p_{h} (respectively p_{v}), and otherwise *closed*, independently of all other edges. Let $P_{\mathbf{p}}$ denote the corresponding probability measure, and let

$$\theta(\mathbf{p}) = P_{\mathbf{p}}(0 \leftrightarrow \infty) \tag{11.114}$$

denote the percolation probability.

One may analyse this model in very much the same way as the homogeneous model on the square lattice, although the lack of full symmetry is a complicating factor. The entire picture will not be presented here, where we restrict ourselves to a computation of the 'critical surface' of the model.

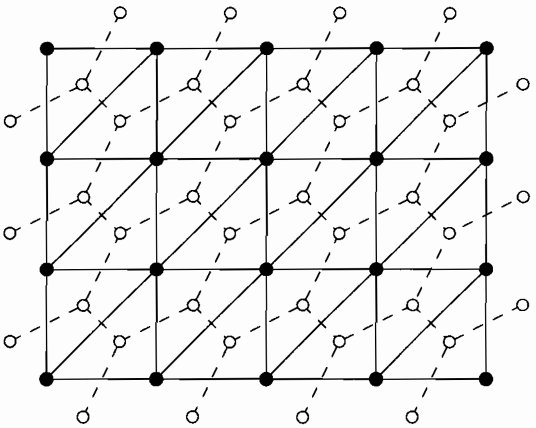


Figure 11.30. The triangular lattice $\mathbb{T}$ and its dual (hexagonal) lattice $\mathbb{H}$. The embedding is unconventional but convenient.

(11.115) Theorem. Critical surface of the inhomogeneous square lattice. *Assume that* $\mathbf{p}$ *is such that* $p_h, p_v < 1$. *We have that*

$$\theta(\mathbf{p}) \begin{cases} = 0 & \text{if } \varphi(\mathbf{p}) \le 1, \\ > 0 & \text{if } \varphi(\mathbf{p}) > 1, \end{cases}$$

where $\varphi(\mathbf{p}) = p_h + p_v$.

The conclusion of this theorem is sometimes stated informally as 'the critical surface is the line $p_h + p_v = 1$'. The theorem generalizes the exact calculation $p_c(\mathbb{L}^2) = \frac{1}{2}$.

We next consider the triangular lattice $\mathbb{T}$, which we embed in the plane in the manner of Figure 11.30. That is, $\mathbb{T}$ is the lattice obtained from $\mathbb{L}^2$ by adding a north-easterly diagonal to each fundamental square. Let $\mathbf{p} = (p_h, p_v, p_d) \in [0, 1]^3$. We declare each horizontal (respectively vertical, diagonal) edge to be *open* with probability p_h (respectively p_v, p_d) and otherwise *closed*, independently of all other edges. Let $P_\mathbf{p}$ denote the corresponding probability measure, and $\theta(\mathbf{p})$ the percolation probability (as in (11.114)).

(11.116) Theorem. Critical surface of the inhomogeneous triangular lattice. *Assume that* $\mathbf{p}$ *is such that* $p_h, p_v, p_d < 1$. *We have that*

$$\theta(\mathbf{p}) \begin{cases} = 0 & \text{if } \psi(\mathbf{p}) \le 1, \\ > 0 & \text{if } \psi(\mathbf{p}) > 1, \end{cases}$$

where $\psi(\mathbf{p}) = p_h + p_v + p_d - p_h p_v p_d$.

Thus the critical surface of the triangular lattice is the set of all $\mathbf{p}$ satisfying the equation $\psi(\mathbf{p}) = 1$. This theorem contains several special cases of interest.

A. *Square lattice.* Set $p_{\mathrm{d}} = 0$ to obtain Theorem (11.115). Although Theorem (11.116) is more general than Theorem (11.115), we shall give a separate proof of the latter theorem below.

B. *Homogeneous triangular lattice.* In the case of the homogeneous model on the triangular lattice, we set $p_{\mathrm{h}} = p_{\mathrm{v}} = p_{\mathrm{d}} = p$, say. The critical value of p satisfies $\psi(p, p, p) = 1$, whence $p_{\mathrm{c}}(\mathbb{T}) = 2\sin(\pi/18)$, in agreement with the statement at the beginning of Section 3.1. We note that the forthcoming proof of Theorem (11.116) may be simplified considerably in this special case, owing to the presence of axes of symmetry.

C. *Inhomogeneous hexagonal lattice.* The triangular lattice $\mathbb{T}$ is dual to the hexagonal lattice $\mathbb{H}$, as drawn in Figure 11.30. Suppose that we call an edge of $\mathbb{H}$ *open* (respectively *closed*) if and only if the corresponding edge of $\mathbb{T}$ is open (respectively closed). Under the product measure $P_{\mathbf{p}}$ given above, the edges of $\mathbb{H}$ are effectively partitioned into three classes, having respective edge-probabilities $p_{\mathrm{h}}, p_{\mathrm{v}}, p_{\mathrm{d}}$. We shall see in the proof of Theorem (11.116) that this model is supercritical if and only if the closed edges of $\mathbb{T}$ constitute a subcritical percolation model, which is to say if $\psi(1 - \mathbf{p}) < 1$ where $1 - \mathbf{p} = (1 - p_{\mathrm{h}}, 1 - p_{\mathrm{v}}, 1 - p_{\mathrm{d}})$. The critical surface is given therefore by the equation $\psi(1 - \mathbf{p}) = 1$. For the homogeneous case, when $p_{\mathrm{h}} = p_{\mathrm{v}} = p_{\mathrm{d}} = p$, the critical probability of $\mathbb{H}$ satisfies

(11.117) $$p_{\mathrm{c}}(\mathbb{H}) = 1 - p_{\mathrm{c}}(\mathbb{T}) = 1 - 2\sin(\pi/18).$$

Further discussion of the relationship between $\mathbb{T}$ and $\mathbb{H}$, and especially of the star–triangle transformation, may be found in the proof of Theorem (11.116).

Proof of Theorem (11.115). This is a straightforward extension of the proof that $p_{\mathrm{c}}(\mathbb{L}^2) = \frac{1}{2}$, and hinges on the following property of self-duality. As usual, we declare each edge of the dual lattice $\mathbb{L}^2_{\mathrm{d}}$ of $\mathbb{L}^2$ to be *open* (respectively *closed*) if it traverses an open edge (respectively closed edge) of $\mathbb{L}^2$. Thus every horizontal (respectively vertical) edge of $\mathbb{L}^2_{\mathrm{d}}$ is open with probability p_{v} (respectively p_{h}). There is a convenient way of harnessing this self-duality, as follows. Let n be a positive integer, and let $D(n)$ be the subgraph of $\mathbb{L}^2$ induced by the set of all vertices within the 'offset diamond' $\{(x_1, x_2) \in \mathbb{R}^2 : |x_1| + |x_2 - \frac{1}{2}| \le n + \frac{1}{2}\}$. We illustrate $D(n)$ in Figure 11.31, together with a certain subgraph $D(n)_{\mathrm{d}}$ of the dual lattice. Note that $D(n)$ and $D(n)_{\mathrm{d}}$ are isomorphic graphs. We shall work henceforth with the pair $D(n)$, $D(n)_{\mathrm{d}}$, although we note that the following arguments may be adapted to rectangles having the more usual orientation. The first step is to note that the proof of Theorem (8.1) may be adapted in order to obtain:

(11.118) either $P_{\mathbf{p}}(I = 0) = 1$ or $P_{\mathbf{p}}(I = 1) = 1$, when $0 < p_{\mathrm{h}}, p_{\mathrm{v}} < 1$,

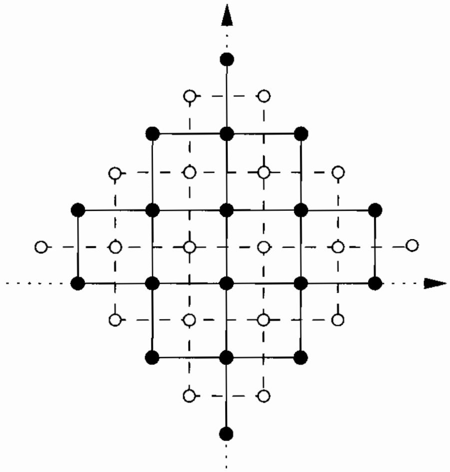


Figure 11.31. The subgraph $D(n)$ of $\mathbb{L}^2$ when $n = 2$, and the associated 'dual' graph $D(n)_d$ of $\mathbb{L}^2_d$.

where I is the number of infinite open clusters of $\mathbb{L}^2$.

Next we prove that

(11.119) $$\theta(\mathbf{p}) = 0 \quad \text{if} \quad \varphi(\mathbf{p}) = 1 \text{ and } p_h, p_v < 1,$$

and we achieve this in the manner of the proof of Lemma (11.12), but working with $D(n)$ in place of the box $[0, n]^2$. We do not give the details of this, noting only that it utilizes (11.118) together with the symmetry of the model under reflections in either the vertical axis of $\mathbb{R}^2$ or in the line $\{(x_1, x_2) \in \mathbb{R}^2 : x_2 = \frac{1}{2}\}$.

Finally, we prove that

(11.120) $$\theta(\mathbf{p}) > 0 \quad \text{if} \quad \varphi(\mathbf{p}) > 1.$$

Assume, on the contrary, that $\theta(\mathbf{r}) = 0$ for some $\mathbf{r} = (r_h, r_v)$ satisfying $\varphi(\mathbf{r}) > 1$. Set

$$\mathbf{p} = (p_h, p_v) = \left(\frac{r_h}{r_h + r_v}, \frac{r_v}{r_h + r_v} \right),$$

so that $\varphi(\mathbf{p}) = 1$ and $p_h < r_h$, $p_v < r_v$. Since $\theta(\mathbf{r}) = 0$, the point (p_h, p_v) lies 'strictly' within the subcritical phase. The proof of Theorem (5.4) may easily be adapted to the current situation, and yields the existence of $\rho = \rho(\mathbf{p}) > 0$ such that

(11.121) $$P_{\mathbf{p}}\big(0 \leftrightarrow \partial S(n)\big) \le e^{-n\rho} \quad \text{for all } n \ge 1,$$

where $S(n) = \{x \in \mathbb{Z}^2 : \delta(0, x) \le n\}$. Note that the proof of Theorem (5.4) relies on no assumption of reflection or rotation symmetry of the percolation model under study.

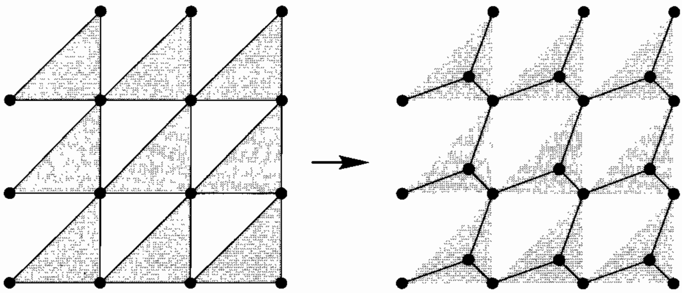


Figure 11.32. In the star–triangle transformation, alternate triangles are replaced by stars.

We now argue as in (11.20) to obtain that

$$P_{\mathbf{p}}(B_n) + P_{\mathbf{p}}(C_n) = 1, \tag{11.122}$$

where B_n (respectively C_n) is the event that $D(n)$ (respectively $D(n)_{\mathrm{d}}$) contains an open path (respectively closed path) joining some vertex on its upper left side to some vertex on its lower right side (respectively some vertex on its lower left side to some vertex on its upper right side). The set of closed edges on $\mathbb{L}^2_{\mathrm{d}}$ is governed by the probability measure $P_{\mathbf{p}'}$ where $\mathbf{p}' = (1 - p_{\mathrm{v}}, 1 - p_{\mathrm{h}}) = (p_{\mathrm{h}}, p_{\mathrm{v}}) = \mathbf{p}$, and therefore

$$P_{\mathbf{p}}(B_n), P_{\mathbf{p}}(C_n) \le (n+1)P_{\mathbf{p}}\big(0 \leftrightarrow \partial S(n)\big) \le (n+1)e^{-n\rho}, \tag{11.123}$$

by (11.121), in contradiction of (11.122) for large n. □

Proof of Theorem (11.116). There are two relations between the triangular and the hexagonal lattices, and this proof uses both of these. The first is duality, and the second is the following 'star–triangle transformation'. Figure 11.32 contains a drawing of the triangular lattice with alternate triangles shaded. Each such shaded triangle may be replaced by a star, as indicated. When every shaded triangle has been thus replaced, the result is the hexagonal lattice.

Next, we add probabilities to this transformation. Consider bond percolation on $\mathbb{T}$ having edge-probabilities p_{h}, p_{v}, p_{d}, as indicated in Figure 11.33. Let T be a typical shaded triangle having corners A, B, C. We may compute the probabilities associated with any open connections joining A, B, and C within the triangle T, namely:

(11.124)

$$\begin{aligned}
P_{\mathbf{p}}\big(A \nleftrightarrow B \text{ and } B \nleftrightarrow C \text{ in } T\big) &= (1-p_{\mathrm{h}})(1-p_{\mathrm{v}})(1-p_{\mathrm{d}}),\\
P_{\mathbf{p}}\big(A \leftrightarrow B \text{ and } B \nleftrightarrow C \text{ in } T\big) &= p_{\mathrm{h}}(1-p_{\mathrm{v}})(1-p_{\mathrm{d}}),\\
P_{\mathbf{p}}\big(A \leftrightarrow B \leftrightarrow C \text{ in } T\big) &= p_{\mathrm{h}}p_{\mathrm{v}}p_{\mathrm{d}} + p_{\mathrm{h}}p_{\mathrm{v}}(1-p_{\mathrm{d}})\\
&\quad + p_{\mathrm{h}}p_{\mathrm{d}}(1-p_{\mathrm{v}}) + p_{\mathrm{v}}p_{\mathrm{d}}(1-p_{\mathrm{h}}),
\end{aligned}$$

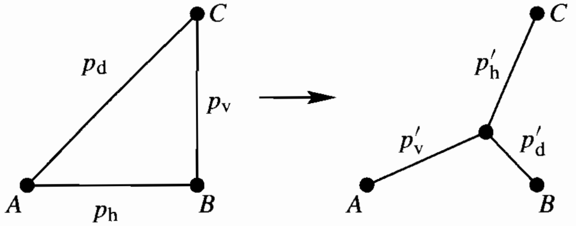


Figure 11.33. Bond probabilities in the star–triangle transformation. The triangle T on the left is replaced by the star S on the right.

and so on. Note that the events in question concern the existence (or not) of open paths within T only.

We now replace the triangle T by the star S, which we think of as being part of the hexagonal lattice $\mathbb{H}$. On $\mathbb{H}$, we consider bond percolation with edge-probabilities p'_{h}, p'_{v}, p'_{d} as illustrated in Figure 11.33. Performing calculations similar to those above, we obtain that

(11.125)

$$\begin{aligned}
P_{\mathbf{p}'}\big(A \not\leftrightarrow B \text{ and } B \not\leftrightarrow C \text{ in } S\big)& \\
&= (1-p'_{\mathrm{h}})(1-p'_{\mathrm{v}})(1-p'_{\mathrm{d}}) + (1-p'_{\mathrm{h}})(1-p'_{\mathrm{v}})p'_{\mathrm{d}} \\
&\quad + (1-p'_{\mathrm{h}})(1-p'_{\mathrm{d}})p'_{\mathrm{v}} + (1-p'_{\mathrm{v}})(1-p'_{\mathrm{d}})p'_{\mathrm{h}}, \\
P_{\mathbf{p}'}\big(A \leftrightarrow B \text{ and } B \not\leftrightarrow C \text{ in } S\big) &= (1-p'_{\mathrm{h}})p'_{\mathrm{v}}p'_{\mathrm{d}}, \\
P_{\mathbf{p}'}\big(A \leftrightarrow B \leftrightarrow C \text{ in } S\big) &= p'_{\mathrm{h}}p'_{\mathrm{v}}p'_{\mathrm{d}},
\end{aligned}$$

and so on. As in (11.124), the events in question relate to the existence (or not) of open paths within S only.

It is a notable fact that the probabilities in (11.124) and (11.125) are identical if the vectors $\mathbf{p}$ and $\mathbf{p}'$ are appropriately chosen, namely if

$$\psi(\mathbf{p}) = 1 \tag{11.126}$$

where $\psi(\mathbf{p}) = p_{\mathrm{h}} + p_{\mathrm{v}} + p_{\mathrm{d}} - p_{\mathrm{h}}p_{\mathrm{v}}p_{\mathrm{d}}$, and in addition

$$p'_{\mathrm{h}} = 1 - p_{\mathrm{h}}, \quad p'_{\mathrm{v}} = 1 - p_{\mathrm{v}}, \quad p'_{\mathrm{d}} = 1 - p_{\mathrm{d}}. \tag{11.127}$$

We leave this to the reader to check. It follows that, if (11.126) and (11.127) hold, then the replacement of T by S is 'invisible' to the rest of the lattice. In particular, all connectivity functions involving vertices of $\mathbb{Z}^2$ are unchanged. We now replace every shaded triangle in Figure 11.32 by the corresponding star. Writing $\mathbb{T}$ and $\mathbb{H}$ for the ensuing lattices, we deduce in particular that

$$P_{\mathbf{p}}\big(0 \leftrightarrow \partial B(n) \text{ in } \mathbb{T}\big) = P_{\mathbf{p}'}\big(0 \leftrightarrow \partial B(n) \text{ in } \mathbb{H}\big)$$

so long as (11.126) and (11.127) hold. We let $n \to \infty$, and find that

$$\theta(\mathbf{p}) = \theta_{\mathbb{H}}(\mathbf{p}') \tag{11.128}$$

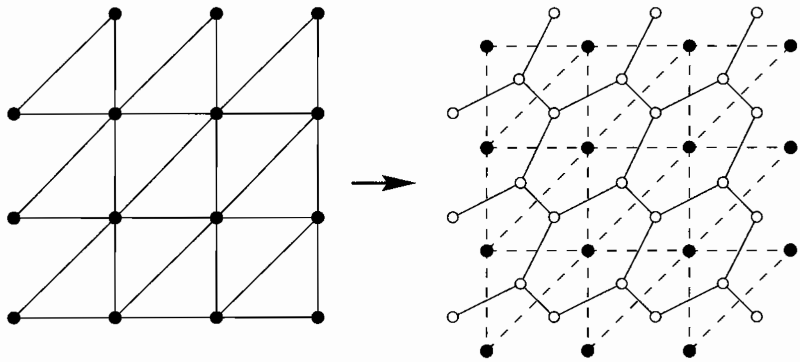


Figure 11.34. The graph $A(n)$ is drawn on the left. The graph on the right with vertices marked by open circles is the graph $H(n)$ obtained from $A(n)$ by the star–triangle transformation; its 'dual' graph $H(n)_d$ has solid vertices and dashed edges. Note that $H(n)_d$ is isomorphic to $A(n)$.

where $\theta_{\mathbb{H}}$ is the percolation probability of the lattice $\mathbb{H}$.

We move now to the proof proper, beginning with the claim that

$$\theta(\mathbf{p}) > 0 \quad \text{if} \quad \psi(\mathbf{p}) > 1. \tag{11.129}$$

This we prove in very much the same way as (11.120). Let n be a positive integer. It is convenient to work with the subgraph $A(n)$ of $\mathbb{T}$ having vertex set $[0, n]^2 \setminus \{(0, n)\}$, and with all induced edges except those joining two vertices $x, y \in A(n)$ satisfying: either $x_1 = y_1 = 0$, or $x_2 = y_2 = n$. The star–triangle transformation maps $A(n)$ onto a graph $H(n)$; the dual of $H(n)$, with a suitable convention about boundary edges, is denoted $H(n)_d$ and is drawn in Figure 11.34.

By the 'top' (respectively 'bottom') of $A(n)$, we mean the set of vertices of $A(n)$ having maximal (respectively minimal) second coordinate; we use a similar terminology for the 'left' and 'right' of $A(n)$, as well as for the graphs $H(n)$ and $H(n)_d$. We abbreviate 'left–right crossing' to 'LR', and 'top–bottom crossing' to 'TB'.

Assume that $\psi(\mathbf{p}) = 1$ and that $\mathbf{p}'$ is given by (11.127). We have by the star–triangle transformation that

$$P_{\mathbf{p}}\big(\text{open TB in } A(n)\big) = P_{\mathbf{p}'}\big(\text{open TB in } H(n)\big),$$

where 'open TB in F' is an abbreviation for the event that F possesses an open top–bottom crossing. Next, we use duality, as in Figure 11.10, to deduce that at least one of the events {open TB in $H(n)$}, {closed LR in $H(n)_d$} must occur; it is possible that both occur. Therefore,

$$P_{\mathbf{p}'}\big(\text{open TB in } H(n)\big) + P_{\mathbf{p}'}\big(\text{closed LR in } H(n)_d\big) \geq 1.$$

Furthermore, by (11.127),

$$P_{\mathbf{p}'}\big(\text{closed LR in } H(n)_{\mathrm{d}}\big) = P_{\mathbf{p}}\big(\text{open LR in } H(n)_{\mathrm{d}}\big),$$

whence

(11.130) $$P_{\mathbf{p}}\big(\text{open TB in } A(n)\big) + P_{\mathbf{p}}\big(\text{open LR in } H(n)_{\mathrm{d}}\big) \geq 1.$$

Assume in contradiction to (11.129) that there exists $\mathbf{r} = (r_{\mathrm{h}}, r_{\mathrm{v}}, r_{\mathrm{d}})$ satisfying $\psi(\mathbf{r}) > 1$ and $\theta(\mathbf{r}) = 0$. Let $\mathbf{p} = \alpha\mathbf{r}$ where $\alpha \in (0, 1)$ is chosen in such a way that $\psi(\mathbf{p}) = 1$. Then $\mathbf{p}$ lies within the subcritical phase of bond percolation on $\mathbb{T}$, and we may use the proof of Theorem (5.4) to deduce as in (11.121) that

$$P_{\mathbf{p}}\big(0 \leftrightarrow \partial S(n)\big) \leq e^{-n\rho} \qquad \text{for } n \geq 1,$$

for some $\rho = \rho(\mathbf{p}) > 0$. As in (11.123), this contradicts (11.130) for large n; therefore (11.129) holds.

The theorem will be proved once we have shown that

(11.131) $$\theta(\mathbf{p}) = 0 \quad \text{if} \quad \psi(\mathbf{p}) = 1 \text{ and } p_{\mathrm{h}}, p_{\mathrm{v}}, p_{\mathrm{d}} < 1.$$

This is rather harder than (11.129) to prove. It seems to be difficult to adapt the argument used earlier for the inhomogeneous square lattice, owing to the lack of an axis of symmetry; instead, we follow an argument of Kesten (1980c). Here is the principal proposition.

(11.132) Proposition. *Consider bond percolation on the triangular lattice* $\mathbb{T}$*, and assume that* $p_{\mathrm{h}}, p_{\mathrm{v}}, p_{\mathrm{d}} < 1$*. If* $\mathbf{p}$ *is such that* $\theta(\mathbf{p}) > 0$*, the origin is almost surely contained in the interior of some open circuit.*

Before proving this, we indicate why it implies (11.131). Assume that $\theta(\mathbf{p}) > 0$, $\psi(\mathbf{p}) = 1$, and $p_{\mathrm{h}}, p_{\mathrm{v}}, p_{\mathrm{d}} < 1$. By Proposition (11.132) and duality, the closed cluster at the origin of the dual lattice $\mathbb{T}_{\mathrm{d}}$ is $P_{\mathbf{p}}$-a.s. finite, which is to say that the percolation probability $\theta_{\mathbb{H}}$ of the hexagonal lattice satisfies $\theta_{\mathbb{H}}(1 - \mathbf{p}) = 0$. However, $\theta_{\mathbb{H}}(1 - \mathbf{p}) = \theta(\mathbf{p})$ by (11.128), in contradiction of the assumption that $\theta(\mathbf{p}) > 0$.

We assume henceforth that $\mathbf{p}$ satisfies

(11.133) $$p_{\mathrm{h}}, p_{\mathrm{v}}, p_{\mathrm{d}} < 1.$$

We shall prove the proposition by building 'half-circuits', and using the fact that inhomogeneous percolation on $\mathbb{T}$ is invariant under reflection in the origin, that is, under the map $x \mapsto -x$ for $x \in \mathbb{Z}^2$.

Here is some notation. Let k (≥ 1) and l be integers. Denote by e_l the vertex $(0, l)$, and by B and B' the boxes $B = B(k)$, $B' = (0, 2l) + B$. Let $C(e_l)$ be the open cluster at e_l, and let

$$C_k(e_l) = \big\{x \in \mathbb{Z}^2 : e_l \leftrightarrow x \text{ off } B \cup B'\big\}.$$

We write $A^{\mathrm{r}}(k, l)$ (respectively $A^{\mathrm{l}}(k, l)$) for the event that $C_k(e_l)$ is unbounded to the right (respectively left); that is, $A^{\mathrm{r}}(k, l)$ is the event that, for all $m \geq 0$, $C_k(e_l)$ contains some vertex whose first coordinate exceeds m, with a similar definition for $A^{\mathrm{l}}(k, l)$.

(11.134) Lemma. *Assume that* $\mathbf{p}$ *satisfies* (11.133) *and* $\theta(\mathbf{p}) > 0$. *Let* $k \ge 1$. *There exists a positive integer* $L = L(k)$ *such that*

$$\max\big\{P_{\mathbf{p}}\big(A^{\mathrm{r}}(k,L)\big),\ P_{\mathbf{p}}\big(A^{\mathrm{l}}(k,L)\big)\big\} \ge \tfrac{1}{3}\theta(\mathbf{p}).$$

Proof. We call $C(e_l)$ *horizontally bounded* if there exists r such that $C(e_l) \subseteq \{x \in \mathbb{R}^2 : |x_1| \le r\}$. Since

$$P_{\mathbf{p}}\big(|C(e_l)| = \infty,\ C(e_l) \text{ is horizontally bounded}\big) = 0,$$

we have that

$$\theta(p) = P_{\mathbf{p}}\big(A^{\mathrm{r}}(l) \cup A^{\mathrm{l}}(l)\big) \le P_{\mathbf{p}}(A^{\mathrm{r}}(l)) + P_{\mathbf{p}}(A^{\mathrm{l}}(l))$$

where $A^{\mathrm{r}}(l)$ (respectively $A^{\mathrm{l}}(l)$) is the event that $C(e_l)$ is unbounded to the right (respectively to the left). Therefore,

$$\max\big\{P_{\mathbf{p}}(A^{\mathrm{r}}(l)),\ P_{\mathbf{p}}(A^{\mathrm{l}}(l))\big\} \ge \tfrac{1}{2}\theta(\mathbf{p}).$$

We suppose henceforth that

(11.135) $$P_{\mathbf{p}}(A^{\mathrm{r}}(l)) \ge \tfrac{1}{2}\theta(\mathbf{p}),$$

otherwise replacing 'right' by 'left' in the following.

We may find an integer $K = K(k)$ and a cylinder event $E = E(k,l)$, defined in terms of the states of edges within the box $e_l + B(K)$, such that

(11.136) $$P_{\mathbf{p}}\big(A^{\mathrm{r}}(l) \,\triangle\, E\big) \le \tfrac{1}{12}\theta(\mathbf{p}) P_{\mathbf{p}}(B \cup B' \text{ closed}),$$

where $\{B \cup B' \text{ closed}\}$ is the event that all edges intersecting $B \cup B'$ are closed. [Note that the right side of (11.136) is strictly positive, by (11.133).] If $l > K+k$, the boxes B, B', $e_l + B(K)$ are disjoint, whence E is independent of the event $\{B \cup B' \text{ closed}\}$. Therefore, if $l > K + k$,

$$\begin{aligned}
P_{\mathbf{p}}\big(A^{\mathrm{r}}(k,l)\big) &= P_{\mathbf{p}}\big(A^{\mathrm{r}}(l) \,\big|\, B \cup B' \text{ closed}\big) \\
&= \frac{P_{\mathbf{p}}(A^{\mathrm{r}}(l),\ B \cup B' \text{ closed})}{P_{\mathbf{p}}(B \cup B' \text{ closed})} \\
&\ge \frac{P_{\mathbf{p}}(E,\ B \cup B' \text{ closed})}{P_{\mathbf{p}}(B \cup B' \text{ closed})} - \tfrac{1}{12}\theta(\mathbf{p}) \\
&= P_{\mathbf{p}}(E) - \tfrac{1}{12}\theta(\mathbf{p}) \\
&\ge P_{\mathbf{p}}(A^{\mathrm{r}}(l)) - \tfrac{1}{6}\theta(\mathbf{p}) \\
&\ge \tfrac{1}{3}\theta(\mathbf{p})
\end{aligned}$$

by (11.135)–(11.136). □

Henceforth we assume that $\theta(\mathbf{p}) > 0$ and that $L = L(k) \geq 1$ satisfies

$$P_{\mathbf{p}}\big(A^{\mathrm{r}}(k, L)\big) \geq \tfrac{1}{3}\theta(\mathbf{p}). \tag{11.137}$$

If this fails, we replace 'right' by 'left' in the following. Having established (11.137), we show next the existence (on the event $A^{\mathrm{r}}(k, L)$) of certain open paths from e_L. For a positive integer n, let H_n denote the line $\{x \in \mathbb{R}^2 : x_1 = n\}$, and let $\mathbb{E}_n$ be the set of all edges of $\mathbb{T}$ having at least one endvertex lying strictly to the left of H_n. Let $V_n(k, l)$ be the number of vertices of H_n which may be reached from e_l by open paths of $\mathbb{E}_n$ which do not intersect $B \cup B'$.

(11.138) Lemma. *Assume that* $\mathbf{p}$ *satisfies* (11.133) *and* $\theta(\mathbf{p}) > 0$. *Let* $k \geq 1$. *There exists* $N = N(k)$ *such that*

$$P_{\mathbf{p}}\big(V_N(k, L) \leq 2L \,\big|\, V_N(k, L) > 0\big) \leq \tfrac{1}{20}. \tag{11.139}$$

Proof. The proof resembles part of that of Lemma (7.9) and inequality (7.41). As in (7.15)–(7.16),

$$\frac{P_{\mathbf{p}}(1 \leq V_n(k, L) \leq 2L)}{P_{\mathbf{p}}(V_n(k, L) > 0)} \leq \frac{\beta^{2L} P_{\mathbf{p}}\big(V_{n+1}(k, L) = 0,\ V_n(k, L) \geq 1\big)}{P_{\mathbf{p}}(V_n(k, L) > 0)}$$

where

$$\beta = \frac{1}{(1 - p_{\mathrm{v}})^2(1 - p_{\mathrm{h}})(1 - p_{\mathrm{d}})} < \infty$$

by (11.133). The numerator tends to 0 as $n \to \infty$, and the denominator approaches $P_{\mathbf{p}}(A^{\mathrm{r}}(k, L))$, which is strictly positive by (11.137). We may now choose n sufficiently large for the claim to hold. □

Here is some more notation. The partial order on $\mathbb{Z}^2$ given by

$$x \leq y \text{ if and only if } x_1 \leq x_2 \text{ and } y_1 \leq y_2$$

induces a total order on H_n which we denote also by '$\leq$'. For $x, y \in H_n$ with $x \leq y$, we write $[x, y]$ for the 'interval' of H_n between x and y, that is,

$$[x, y] = \{z \in H_n : x \leq z \leq y\},$$

with a similar definition for open, half-open, and unbounded intervals of H_n.

Given an interval I of H_n, we say that e_l *is joined to* I if there exists an open path from e_l to some vertex of I using only edges of $\mathbb{E}_n$ but not intersecting B. Our target is to find four ordered disjoint intervals I_1, I_2, I_3, I_4 of H_n such that e_L is joined to I_1 and to I_3, and e_{-L} is joined to I_2 and to I_4; see Figure 11.35 for an illustration of the required events. Since $\mathbb{T}$ is planar, this will imply that

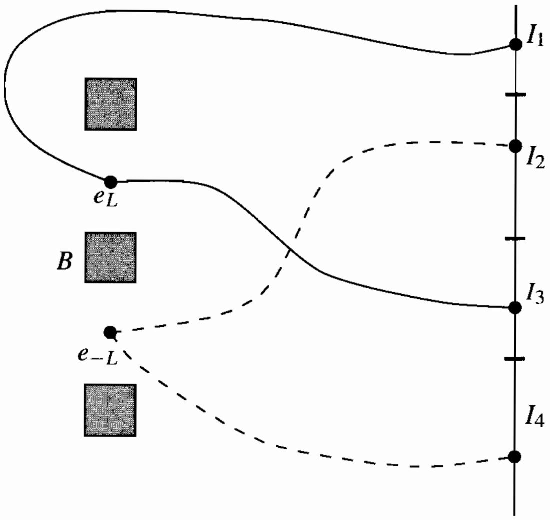


Figure 11.35. If e_L is joined to I_1 and to I_3, and e_{-L} is joined to I_2 and to I_4, then $e_L \leftrightarrow e_{-L}$.

e_L and e_{-L} are connected by an open path off B. The model is invariant under reflection in the origin, and this will enable a proof via the FKG inequality that B is surrounded (with probability bounded away from 0) by an open path.

We have that $\theta(\mathbf{p}) > 0$, and we let $k \geq 1$, and let L and N be given according to (11.137) and Lemma (11.138). There will be two cases to consider, depending on whether or not

(11.140) $$P_{\mathbf{p}}\big(V^+ < \infty \,\big|\, V^+ > 0\big) \geq \tfrac{1}{2},$$

where $V^+ = V_N(k, L)$.

On the event $\{V^+ < \infty\}$, we denote by $c_1^+, c_2^+, \dots, c_{V^+}^+$ the vertices on H_n which contribute to V^+; we write the c_i^+ in decreasing order, so that $c_i^+ > c_{i+1}^+$ for all i. In order to control the positions of c_1^+ and $c_{V^+}^+$, we introduce some 'percentiles'. For $0 < \alpha < 1$, let y_α and z_α be vertices on H_n satisfying:

(11.141)
$$P_{\mathbf{p}}\big(c_1^+ > y_\alpha \,\big|\, 0 < V^+ < \infty\big) < \alpha \leq P_{\mathbf{p}}\big(c_1^+ \geq y_\alpha \,\big|\, 0 < V^+ < \infty\big),$$
$$P_{\mathbf{p}}\big(c_{V^+}^+ > z_\alpha \,\big|\, 0 < V^+ < \infty\big) < \alpha \leq P_{\mathbf{p}}\big(c_{V^+}^+ \geq z_\alpha \,\big|\, 0 < V^+ < \infty\big).$$

It is easily seen that y_α and z_α are non-increasing functions of α.

We now turn our attention from the vertex e_L to the vertex e_{-L}. By the translation invariance of $\mathbb{T}$, we have that

(11.142) $$P_{\mathbf{p}}\big(V^- < \infty \,\big|\, V^- > 0\big) = P_{\mathbf{p}}\big(V^+ < \infty \,\big|\, V^+ > 0\big),$$

where $V^- = V_N(k, -L)$. When $V^- < \infty$, we write $c_1^-, c_2^-, \dots, c_{V^-}^-$ for the vertices on H_n which contribute to V^-, in decreasing order as before. It follows from (11.141) and the translation invariance of $\mathbb{T}$ that

(11.143)
$$P_{\mathbf{p}}\big(c_1^- > y_\alpha - (0,2L)\,\big|\,0 < V^- < \infty\big) < \alpha \le P_{\mathbf{p}}\big(c_1^- \ge y_\alpha - (0,2L)\,\big|\,0 < V^- < \infty\big),$$
$$P_{\mathbf{p}}\big(c_{V^-}^- > z_\alpha - (0,2L)\,\big|\,0 < V^- < \infty\big) < \alpha \le P_{\mathbf{p}}\big(c_{V^-}^- \ge z_\alpha - (0,2L)\,\big|\,0 < V^- < \infty\big).$$

We now define intervals I_j of H_n by

$$\begin{aligned} I_1 &= [y_{1/5}, \infty), \\ I_2 &= \big[y_{2/5} - (0,2L), y_{1/5}\big), \\ I_3 &= \big[z_{4/5}, y_{2/5} - (0,2L)\big), \\ I_4 &= (-\infty, z_{4/5}). \end{aligned}$$

It will be important to know that

(11.144) $$z_{3/5} < y_{2/5} - (0,2L) \quad \text{if (11.140) holds;}$$

since $z_{4/5} \le z_{3/5}$, (11.144) implies in particular that the I_j are disjoint and non-empty under (11.140). Relation (11.144) is proved as follows. Assume that (11.140) holds. We have that

$$P_{\mathbf{p}}\big(V^+ \le 2L\,\big|\,0 < V^+ < \infty\big) = \frac{P_{\mathbf{p}}(V^+ \le 2L \mid V^+ > 0)}{P_{\mathbf{p}}(V^+ < \infty \mid V^+ > 0)} \le \tfrac{1}{10}$$

by Lemma (11.138) and (11.140), and therefore,

$$\begin{aligned} P_{\mathbf{p}}\big(c_{V^+}^+ &< y_{2/5} - (0,2L)\,\big|\,0 < V^+ < \infty\big) \\ &\ge P_{\mathbf{p}}\big(c_1^+ \le y_{2/5}\,\big|\,0 < V^+ < \infty\big) - P_{\mathbf{p}}\big(V^+ \le 2L\,\big|\,0 < V^+ < \infty\big) \\ &\ge \tfrac{3}{5} - \tfrac{1}{10} > \tfrac{2}{5}, \end{aligned}$$

which implies (11.144).

(11.145) Lemma. *Assume that* $\mathbf{p}$ *satisfies* (11.133) *and* $\theta(\mathbf{p}) > 0$. *Let* $k \ge 1$, *and let* L *and* N *be given by* (11.137) *and Lemma* (11.138). *It is the case that*

$$P_{\mathbf{p}}\big(e_L \leftrightarrow e_{-L} \text{ off } B\big) \ge \big(\tfrac{1}{30}\theta(\mathbf{p})\big)^4.$$

Proof. There are two cases. Assume first that (11.140) holds, so that (11.144) applies. In this case,

(11.146)
$$\begin{aligned} P_{\mathbf{p}}(e_L \text{ is joined to } I_1) &\ge P_{\mathbf{p}}\big(0 < V^+ < \infty,\ c_1^+ \ge y_{1/5}\big) \\ &\ge \tfrac{1}{5} P_{\mathbf{p}}\big(V^+ < \infty\,\big|\,V^+ > 0\big) P_{\mathbf{p}}(V^+ > 0) \\ &\ge \tfrac{1}{30}\theta(\mathbf{p}) \end{aligned}$$

by (11.140)–(11.141) and (11.137). Similarly,

(11.147)
$$\begin{aligned}P_{\mathbf{p}}(e_{-L} &\text{ is joined to } I_2)\\ &\geq P_{\mathbf{p}}\big(0 < V^- < \infty,\ c_1^- \in I_2\big)\\ &\geq P_{\mathbf{p}}\big(0 < V^- < \infty,\ y_{2/5} - (0, 2L) \leq c_1^- \leq y_{1/5} - (0, 2L)\big)\\ &\geq \big(\tfrac{2}{5} - \tfrac{1}{5}\big) P_{\mathbf{p}}\big(V^- < \infty \,\big|\, V^- > 0\big) P_{\mathbf{p}}(V^- > 0)\\ &\geq \tfrac{1}{30}\theta(\mathbf{p})\end{aligned}$$

by (11.140)–(11.143) and (11.137).

Also, by (11.144),

(11.148)
$$\begin{aligned}P_{\mathbf{p}}(e_L &\text{ is joined to } I_3)\\ &\geq P_{\mathbf{p}}\big(0 < V^+ < \infty,\ c_{V^+}^+ \in I_3\big)\\ &\geq P_{\mathbf{p}}\big(0 < V^+ < \infty,\ z_{4/5} \leq c_{V^+}^+ \leq z_{3/5}\big)\\ &\geq \big(\tfrac{4}{5} - \tfrac{3}{5}\big) P_{\mathbf{p}}\big(V^+ < \infty \,\big|\, V^+ > 0\big) P_{\mathbf{p}}(V^+ > 0)\\ &\geq \tfrac{1}{30}\theta(\mathbf{p}),\end{aligned}$$

with a similar inequality for $P_{\mathbf{p}}(e_{-L}$ is joined to $I_4)$.

By the FKG inequality,

$$P_{\mathbf{p}}\big(e_L \text{ is joined to } I_1 \text{ and } I_3,\, e_{-L} \text{ is joined to } I_2 \text{ and } I_4\big) \geq \big(\tfrac{1}{30}\theta(\mathbf{p})\big)^4.$$

As indicated in Figure 11.35, since $\mathbb{T}$ is planar, such open paths must contain a path joining e_L to e_{-L}. The claim of the lemma follows.

Assume next that (11.140) is false, which is to say that

$$P_{\mathbf{p}}\big(V^+ = \infty \,\big|\, V^+ > 0\big) > \tfrac{1}{2}.$$

In this case the required argument is somewhat simpler. Let V_{b}^+ (respectively V_{t}^+) be the number of vertices $x\ (\in H_n)$ contributing to V^+ having $x_2 < L$ (respectively $x_2 > L$). If $V^+ = \infty$ then either $V_{\mathrm{b}}^+ = \infty$ or $V_{\mathrm{t}}^+ = \infty$, and we shall therefore assume that

(11.149)
$$P_{\mathbf{p}}\big(V_{\mathrm{b}}^+ = \infty \,\big|\, V^+ > 0\big) > \tfrac{1}{4}.$$

If this is false, then we replace V_{b}^+ by V_{t}^+ and make suitable changes in the following. We write V_{b}^- for the number of vertices $x\ (\in H_n)$ contributing to V^- and having $x_2 < -L$, and we have by translation invariance that

(11.150)
$$P_{\mathbf{p}}\big(V_{\mathrm{b}}^- = \infty \,\big|\, V^- > 0\big) > \tfrac{1}{4}.$$

Now, if $V_b^+ = V_b^- = \infty$, there exist $u_1, u_2, u_3, u_4 \in H_n$ satisfying $u_1 > u_2 > u_3 > u_4$, such that e_L is joined to u_1 and u_3, and e_{-L} is joined to u_2 and u_4. It follows as before that

$$(11.151) \qquad P_{\mathbf{p}}(e_L \leftrightarrow e_{-L} \text{ off } B) \geq P_{\mathbf{p}}(V_b^+ = V_b^- = \infty).$$

By the FKG inequality, (11.149)–(11.150), and (11.137),

$$\begin{aligned} P_{\mathbf{p}}(V_b^+ = V_b^- = \infty) &\geq P_{\mathbf{p}}(V_b^+ = \infty) P_{\mathbf{p}}(V_b^- = \infty) \\ &> \tfrac{1}{16} P_{\mathbf{p}}(V^+ > 0) P_{\mathbf{p}}(V^- > 0) \\ &\geq \tfrac{1}{16} \big(\tfrac{1}{3}\theta(\mathbf{p})\big)^2, \end{aligned}$$

whence the conclusion of the lemma holds. □

There are two straightforward steps in deducing Proposition (11.132) from Lemma (11.145). Let α be a path of $\mathbb{T}$ joining e_L to e_{-L} which does not intersect B, and let $\Delta(\alpha)$ be the change in argument of α as we move from e_L to e_{-L}. Clearly, $\Delta(\alpha) = 2\pi\delta(\alpha) + \pi$ for some integer $\delta(\alpha)$. We write $e_L \overset{+}{\longleftrightarrow} e_{-L}$ (respectively $e_L \overset{-}{\longleftrightarrow} e_{-L}$) if there exists an open path α from e_L to e_{-L} not intersecting B and with $\delta(\alpha) \geq 0$ (respectively $\delta(\alpha) < 0$). Since the process is invariant under reflection in the origin, and since reflection in the origin maps paths α with $\delta(\alpha) \geq 0$ to paths α' with $\delta(\alpha') < 0$, we have that

$$P_{\mathbf{p}}(e_L \overset{+}{\longleftrightarrow} e_{-L}) = P_{\mathbf{p}}(e_L \overset{-}{\longleftrightarrow} e_{-L}).$$

Furthermore,

$$\begin{aligned} \big(\tfrac{1}{30}\theta(\mathbf{p})\big)^4 &\leq P_{\mathbf{p}}\big(e_L \leftrightarrow e_{-L} \text{ off } B\big) \\ &\leq P_{\mathbf{p}}(e_L \overset{+}{\longleftrightarrow} e_{-L}) + P_{\mathbf{p}}(e_L \overset{-}{\longleftrightarrow} e_{-L}) \end{aligned}$$

by Lemma (11.145). Therefore, by the FKG inequality,

$$\begin{aligned} P_{\mathbf{p}}\big(e_L \overset{+}{\longleftrightarrow} e_{-L}, e_L \overset{-}{\longleftrightarrow} e_{-L}\big) &\geq P_{\mathbf{p}}(e_L \overset{+}{\longleftrightarrow} e_{-L}) P_{\mathbf{p}}(e_L \overset{-}{\longleftrightarrow} e_{-L}) \\ &\geq \tfrac{1}{4}\big(\tfrac{1}{30}\theta(\mathbf{p})\big)^8. \end{aligned}$$

However, if $e_L \overset{+}{\longleftrightarrow} e_{-L}$ and $e_L \overset{-}{\longleftrightarrow} e_{-L}$, the union of the two corresponding paths must contain an open circuit having $B = B(k)$ in its interior. It follows that

$$(11.152) \quad P_{\mathbf{p}}\big(B(k) \text{ lies in the interior of some open circuit}\big) \geq \eta \quad \text{for all } k \geq 1.$$

where $\eta = \frac{1}{4}\big(\frac{1}{30}\theta(\mathbf{p})\big)^8$.

Let $A(u, v)$ denote the annulus $B(v) \setminus B(u)$, and let $O(u, v)$ be the event that $A(u, v)$ contains an open circuit having $B(u)$ in its interior. The estimate (11.152) is valid for all k, and we construct a sequence $k_1 < k_2 < k_3 < \cdots$ as follows. Take $k_1 = 1$. Find k_2 such that

$$P_{\mathbf{p}}\big(O(k_1, k_2)\big) \geq \tfrac{1}{2}\eta,$$

and more generally find k_{n+1} such that

$$P_{\mathbf{p}}\big(O(k_n, k_{n+1})\big) \geq \tfrac{1}{2}\eta.$$

The events $O(k_1, k_2), O(k_2, k_3), \ldots$ are independent, each having probability at least $\frac{1}{2}\eta$. At least one of them (actually infinitely many of them) must occur with probability 1, and Proposition (11.132) has been proved. □

11.10 Notes

Section 11.2. Duality is used in independent work of Hammersley (1959) and Harris (1960). See Bondy and Murty (1976, p. 140) for a graph-theoretic reference, and Kesten (1982, p. 386) for a more careful treatment than that presented here. Theorem (11.4) appears in Sykes and Essam (1964) together with related material. Their 'exact calculation' contains a hypothesis which remains unjustified to this day, namely that κ has a singularity at p_c; see the discussion after Theorem (4.31). When applied to other pairs $(\mathcal{L}, \mathcal{L}_d)$ of dual lattices, the Sykes–Essam argument indicates that $p_c(\mathcal{L}) + p_c(\mathcal{L}_d) = 1$. Taken in conjunction with the so-called star–triangle transformation, this implies that the critical probability of bond percolation on the triangular (respectively hexagonal) lattice is $2\sin(\pi/18)$ (respectively $1 - 2\sin(\pi/18)$). These exact values have been justified by rigorous arguments of Wierman (1981). See Chapter 3 and Section 11.9 also.

Section 11.3. Kesten (1980a) proved that $p_c = \frac{1}{2}$, building on work of Harris (1960), Russo (1978), and Seymour and Welsh (1978). Lemma (11.12) was first proved using quite different techniques by Harris (1960) in a remarkable paper which contained several of the techniques used in later developments. Such results were extended to other dual pairs of two-dimensional lattices by Fisher (1961) and Fisher and Essam (1961), who observed also that similar techniques work for site percolation on so called 'matching pairs' of lattices (see Kesten (1982, Section 2.2)). Harris's proof of Lemma (11.13) has been adapted with considerable originality by Gandolfi, Keane, and Russo (1988) to prove the uniqueness of the infinite open cluster for a certain class of 'non-Bernoulli' percolation processes. Related arguments appear in the proof of Theorem (11.116).

The simple proof given here that $\theta(\frac{1}{2}) = 0$ is due to Zhang (1988). The geometrical ingredients of our proofs that $p_c \leq \frac{1}{2}$ are contained in Russo (1978) and Seymour and Welsh (1978).

The arguments of this section may be adapted to show with the aid of the star–triangle transformation that the critical probabilities of the triangular and hexagonal lattices are as given above. These arguments may be applied also to certain inhomogeneous processes. See Section 11.9.

We saw in Lemma (11.21) that there is probability $\frac{1}{2}$ of a left–right crossing of the rectangle $[0, n+1] \times [0, n]$ when $p = p_c$. There is a very remarkable proposal known as 'Cardy's formula' (after Cardy (1992)) which appears to give the exact limiting values of all crossing probabilities of large regions, for all critical percolation models in two dimensions.

Consider a critical percolation model in two dimensions. Let γ be a simple closed curve, and let z_1, z_2, z_3, z_4 be four points on γ in clockwise order. We denote by α the segment of γ from z_1 to z_2, and by β the segment from z_3 to z_4. We write $G_r(\gamma, \alpha, \beta)$ for the event that there exists an open path joining $r\alpha$ to $r\beta$ using edges inside $r\gamma$. [Here, $r\gamma$, $r\alpha$, $r\beta$ are the dilations of γ, α, β by the positive factor r.] It is believed that the limit

$$\pi(\gamma, \alpha, \beta) = \lim_{r\to\infty} P_{p_c}\big(G_r(\gamma, \alpha, \beta)\big)$$

exists for all triples (γ, α, β).

Now, there exists a conformal map φ on the interior of γ which maps it to the unit disc, taking γ to its circumference S, and the points z_i to some points w_i. There are many such maps, but the cross-ratio of such maps,

$$u = \frac{(w_4 - w_3)(w_2 - w_1)}{(w_3 - w_1)(w_4 - w_2)},$$

is a constant satisfying $0 \le u \le 1$ (we think of z_i and w_i as points in the complex plane $\mathbb{C}$). We may parametrize the w_i as follows: we may assume that

$$w_1 = e^{i\theta}, \quad w_2 = e^{-i\theta}, \quad w_3 = -e^{i\theta}, \quad w_4 = -e^{-i\theta},$$

for some θ satisfying $0 \le \theta \le \pi/2$. Note that $u = \sin^2\theta$.

The proposal of conformal invariance implies that the function π satisfies $\pi(\gamma, \alpha, \beta) = \pi(S, \varphi(\alpha), \varphi(\beta))$, and that $f(u) = \pi(S, \varphi(\alpha), \varphi(\beta))$ satisfies the differential equation

$$u(1-u)f''(u) + \tfrac{2}{3}(1-2u)f'(u) = 0,$$

subject to the boundary conditions $f(0) = 0$, $f(1) = 1$. The solution is the hypergeometric function

$$f(u) = \frac{3\Gamma(\frac{2}{3})}{\Gamma(\frac{1}{3})^2} u^{1/3} {}_2F_1(\tfrac{1}{3}, \tfrac{2}{3}, \tfrac{4}{3}; u).$$

This prediction for the values of crossing probabilities is known as 'Cardy's formula'.

The derivation is somewhat speculative, but lengthy computer calculations support the formula; see Langlands, Pichet, Pouliot, and Saint-Aubin (1992) and Langlands, Pouliot, and Saint-Aubin (1994).

Another 'exact formula' for crossing probabilities has been proposed by Watts (1996).

Section 11.4. Theorem (11.24) appears in Chayes, Chayes, Grimmett, Kesten, and Schonmann (1989). The argument of the proof of Theorem (11.25) is implicit in Russo (1978). Lemma (11.27) is implicit in the above paper of Chayes et al. (1989) and appears explicitly in Aizenman, Chayes, Chayes, and Newman (1988) and Durrett and Schonmann (1988b).

Alexander, Chayes, and Chayes (1990) have studied the typical shape of large finite open clusters in two dimensions, thereby verifying the so called 'Wulff construction'. In so doing, they have proved the existence of the limit of

$$\frac{1}{\sqrt{n}} \log P_p(|C| = n)$$

when $p > \frac{1}{2}$ and $d = 2$. See Cerf (1998a) for a large deviation principle associated with such clusters, and Cerf (1998b) for similar results in three dimensions.

Section 11.5. Most of this material appears in Grimmett (1981a, 1983); we owe to Chayes and Chayes (1986b) the observation that the function f may be picked in such a way that there exists an infinite open cluster *at* the critical point. The latter authors study the corresponding problem in three and more dimensions also, obtaining improvements to results of Hammersley and Whittington (1985) and Grimmett (1985a). For corresponding results for models of ferromagnetism, see Chayes and Chayes (1986b) and Aizenman, Chayes, Chayes, and Newman (1988).

In related work, Grimmett (1981a) studied the quantity

$$M_n = \max\big\{m : \text{there exists a left–right crossing of the rectangle } [0, m] \times [0, n]\big\},$$

showing that

$$\text{if } p < \tfrac{1}{2} \text{ then } M_n/\log n \to \xi(p) \text{ in probability,}$$
$$\text{if } p > \tfrac{1}{2} \text{ then } n^{-1} \log M_n \to \xi(1-p)^{-1} \text{ in probability,}$$

as $n \to \infty$. Durrett and Schonmann (1988b) have proved that $M_n/E_p(M_n)$ has a limit distribution which is the exponential distribution with mean 1 when $\frac{1}{2} < p < 1$.

Section 11.6. The basic references for central limit theorems are Brånvall (1980), Cox and Grimmett (1981, 1984), Herrndorf (1985), and Zhang (1998). As for results about large deviations, Durrett and Schonmann (1988a) have studied the probabilities of large deviations in the number of vertices inside $B(n)$ which are contained in the infinite open cluster. Many of the arguments used to prove central limit theorems in two dimensions are applicable in higher dimensions also.

Section 11.7. The RSW theorem appears in Russo (1978, 1981) and Seymour and Welsh (1978). Chayes and Chayes (1986a) describe interesting and useful constructions in the spirit of Lemma (11.75).

Section 11.8. Theorems (11.89) and (11.93) appear with related material in Kesten (1981, 1982), where one may find also the proof that κ is twice differentiable at $\frac{1}{2}$. The exception is the lower bound for $P_{1/2}(0 \leftrightarrow \partial B(n))$, and the consequences thereof; this is due to van den Berg and Kesten (1985). Our proofs of the upper bounds for χ and χ^{f} are based on Menshikov's proof of the uniqueness of the critical point (Theorem (5.4)) rather than on Kesten's original argument; the majority of our argument is valid for any number of dimensions, but we lack the starting point '$\rho < \infty$' if $3 \le d < 19$; see the relevant remarks in the notes for Section 10.2.

Section 11.9. Theorem (11.115) was proved by Kesten (1982, p. 54), using other methods than those presented here.

The inhomogeneous triangular lattice presents a special difficulty, since it lacks any axis of symmetry. The proof of Theorem (11.116) was begun in Kesten (1980c, 1982, p. 381) and apparently completed in Kesten (1988a), but full details were never published. The second half of the proof given here is derived from Kesten (1980c). Kesten has proved a more general version of Proposition (11.132), valid for site percolation on general two-dimensional lattices which are part of a matching pair. It is apparently hard to adapt the method of Zhang (1988) to such an inhomogeneous model.

Proposition (11.132) has been generalized by Gandolfi, Keane, and Russo (1988) in the direction of ergodic measures rather than just product measures, but at the price of assuming that the model has axes of symmetry. Their result may not be applied to the inhomogeneous triangular lattice model studied in this section. Such arguments have some of their origins in the beautiful paper of Harris (1960).

It should be noted that Sykes and Essam (1964) 'predicted' the exact forms of the critical surfaces of the inhomogeneous square and triangular lattices, by utilizing duality, the star–triangle transformation, and an argument concerning the number of open clusters per vertex; cf. Theorem (11.4).

Chapter 12

Extensions of Percolation

12.1 Mixed Percolation on a General Lattice

Bond percolation on $\mathbb{L}^d$ is an attractive representative of the class of all percolation processes. It is simple to formulate, and its analysis contains a minimum of geometrical and analytical detail whilst retaining all general features of substance. We may generalize in many directions, of which we mention a few in this section.

A. *Other lattices*. The hypercubic lattice $\mathbb{L}^d$ is a simple and convenient structure but, apart from self-duality in two dimensions, it has few magical properties. With the exception of the results for bond percolation on $\mathbb{L}^2$, essentially all the results of Chapters 1–10 may be reformulated for bond percolation on a general lattice $\mathcal{L}$. Here are two formal definitions of a lattice. Kesten (1982, pp. 10–12) defines a *periodic graph in d dimensions* to be a connected, loopless graph $\mathcal{L}$, with bounded vertex degrees, which is embedded in $\mathbb{R}^d$ in such a way that:

(a) the translations $x \to x + e$ are automorphisms of $\mathcal{L}$ for each unit vector e parallel to a coordinate axis;

(b) all edges are of finite length; and

(c) every compact subset of $\mathbb{R}^d$ intersects only finitely many edges.

Such a definition stresses the geometrical aspects of a lattice. Graph-theoretic aspects are stressed more by Grimmett (1978b) in the following definition. A *d-dimensional lattice* is an infinite locally finite graph which admits as a group $\mathcal{G}$ of automorphisms the product of d cyclic groups acting like $\mathbb{Z}^d$ such that: the cyclic subgroups of $\mathcal{G}$ are generated by automorphisms without fixed points, and the vertex set of $\mathcal{L}$ has only finitely many orbits under $\mathcal{G}$. See also Rogers (1964), Biggs (1976), and Godsil and McKay (1980). Needless to say, most percolation theorists avoid such definitions. We defer until Section 12.3 a discussion of 'long-range' percolation.

B. *Percolation on other graphs.* A useful property of finite-dimensional lattices is that the surface/volume ratio of a finite box B tends to 0 in the limit as B fills out space. A notable family of graphs for which this property does not hold is the set of regular trees. Percolation on a regular tree (such as the binary tree of Section 10.1) differs qualitatively from percolation on $\mathbb{Z}^d$ in that there exist a.s. infinitely many infinite open clusters whenever there exists at least one. This observation led Grimmett and Newman (1990) to study percolation on the product $T \times \mathbb{Z}$ of a regular tree T and the integers $\mathbb{Z}$. This percolation process has three phases, characterized by whether the number of infinite open clusters equals (a.s.) 0, ∞, or 1.

Later authors have investigated such phenomena in detail in the context of percolation on general graphs and on Cayley graphs of general groups, and a rich structure has emerged. See Benjamini (1996), Benjamini and Schramm (1996), Benjamini, Lyons, Peres, and Schramm (1997, 1998), Häggström and Peres (1998), and Schonmann (1998a, b).

C. *Site percolation.* Suppose that we wish to block the *vertices* of a lattice rather than its edges. We may declare each vertex to be *open* with probability p and *closed* otherwise, independently of all other vertices. The open clusters of this 'site percolation' model are the connected components of the subgraph of the lattice induced by the set of open vertices. In broad terms, all the results of Chapters 1–10 have analogues for site percolation processes.

We saw in Section 1.6 that site percolation is more general than bond percolation in the following sense: every bond process may be reformulated as a site process on a different lattice, but there exist site processes which do not arise thus from bond processes.

D. *Mixed percolation.* In mixed percolation models, both edges and vertices may be open or closed, possibly with different probabilities. See McDiarmid (1980, 1981) and Hammersley (1980). One may even block the faces, in two dimensions at least; see Wierman (1984b).

E. *Inhomogeneous percolation.* Here are two examples of inhomogeneous bond percolation processes:

(i) bond percolation on $\mathbb{L}^2$ in which each horizontal edge is open with probability p_h, and each vertical edge is open with probability p_v;

(ii) bond percolation on the triangular lattice in which an edge is open with probability p_1, p_2, or p_3, depending on whether its inclination to the horizontal is 0, $\pi/3$, or $2\pi/3$.

We refer the reader to Section 11.9 and to Kesten (1982, Chapter 3) for formal definitions and some relevant technology. Inhomogeneous processes have properties broadly similar to those of homogeneous processes, although the critical probability p_c of the latter becomes the 'critical surface' $\varphi(p_1, p_2, \ldots) = 0$ of the former, where $p_1, p_2, \ldots$ are the edge-probabilities of the process in question. We make this remark more transparent by reference to the two examples above.

(i) For the square lattice we define $\varphi(p_h, p_v) = p_h + p_v - 1$. It turns out that all open clusters are almost surely finite if $\varphi(p_h, p_v) < 0$, and there exists (almost surely) a unique infinite open cluster if $\varphi(p_h, p_v) > 0$.

(ii) The same conclusions are valid in the case of the triangular lattice with the function φ defined by $\varphi(p_1, p_2, p_3) = p_1 + p_2 + p_3 - p_1 p_2 p_3 - 1$.

These two statements are proved in Section 11.9, and their history is summarized in the notes for that section.

12.2 *AB* Percolation

We are provided with a lattice $\mathcal{L}$ and two types of atom, type A and type B. Each vertex of $\mathcal{L}$ is occupied by an atom of one or other type, and we suppose that there is probability p that any given vertex is occupied by an atom of type A. We assume further that different vertices are occupied independently of each other. Each edge of $\mathcal{L}$ is declared to be *open* if its endvertices are occupied by atoms of *different* types; the idea is that dissimilar atoms bond together, whereas similar atoms repel each other. Of principal interest is the probability $\theta(p; \mathcal{L})$ that the origin of the lattice belongs to an infinite cluster of open edges; we may think of $\theta(p; \mathcal{L})$ as being the probability that the origin is an endvertex of an infinite 'AB path'.

As a model for gelation, this process has been studied somewhat intuitively by Halley (1983) and, under the title 'antipercolation', by Sevsek, Debierre, and Turban (1983). Here is a pencil sketch of results of Appel and Wierman (1987), Scheinerman and Wierman (1987), and Wierman and Appel (1987). In the negative direction, it is known that $\theta(p; \mathcal{L}) = 0$ for all values of p if $\mathcal{L}$ belongs to a certain class of bipartite lattices; the square lattice does not lie in this general class, but separate arguments may be used to establish $\theta(p; \mathbb{L}^2) = 0$ for all p. In the positive direction, we have that $\theta(p; \mathcal{L}) > 0$ if $\mathcal{L}$ is the triangular lattice, for values of p lying in a certain non-empty interval containing $\frac{1}{2}$. We note that currently known results have most power when applied to two-dimensional lattices.

For further work, see Łuczak and Wierman (1989), Wierman (1989), Appel and Wierman (1993), and Benjamini and Kesten (1995).

12.3 Long-Range Percolation in One Dimension

Bond percolation on $\mathbb{L}^d$ is a process which involves 'nearest neighbour interaction' only, in the sense that only pairs of neighbours may be joined by open edges. Greater generality is obtained by considering 'long-range' processes, constructed as follows. Let $\mathbf{p} = (p(x) : x \in \mathbb{Z}^d)$ be a collection of numbers from the interval $[0, 1)$, indexed by the set $\mathbb{Z}^d$ of d-vectors of integers; we suppose that $p(x) = p(-x)$ for all $x \in \mathbb{Z}^d$. We examine in turn each unordered pair $\{u, v\}$ of distinct points in $\mathbb{Z}^d$; we join u and v by an open edge with probability $p(v - u)$, independently of all other pairs. The resulting process is a translation-invariant long-range bond percolation process, and we may recover ordinary bond percolation with edge-probability p by setting

$$p(x) = \begin{cases} p & \text{if } x \text{ or } -x \text{ is a unit vector,} \\ 0 & \text{otherwise.} \end{cases}$$

We have ruled out the possibility that $p(x) = 1$ in order to avoid certain cases of little intrinsic interest.

We define the *range* of a long-range process to be the integer M given by

$$M = \inf\{m : p(x) = 0 \text{ if } |x| > m\},$$

with the convention that the infimum of the empty set is ∞. We call the process *finite-range* if $M < \infty$ and *infinite-range* otherwise. A finite-range process with range M may be reformulated as a nearest-neighbour process on the lattice obtained by decorating $\mathbb{Z}^d$ with edges between all pairs of points which are within distance M of each other. The ensuing process will generally be 'inhomogeneous', since different edges may have different probabilities of being open, but such inhomogeneity is unlikely to affect substantially the qualitative features of the process over long distances. On the other hand, the consequences of *infinite-range* interactions are somewhat less predictable, and it is upon such processes that we concentrate here. [We recall, in addition, the beautiful results of Hara and Slade (1990, 1994) for 'spread out' models in 7 and more dimensions. They have proved mean field behaviour for such systems; see Section 10.3.]

It is perhaps in one dimension that the effect of infinite-range interactions is most startling. Whereas nearest-neighbour percolation on the line $\mathbb{L}$ has few (or no) properties of interest to mathematical physicists, infinite-range processes have rich and beautiful structure. This will not surprise those familiar with the analogous one-dimensional Ising model, of which we summarize briefly the properties. Let us consider the Ising model of ferromagnetism on the line $\mathbb{Z}$, in which each pair $\{u, v\}$ of points has an interaction of strength $J(|v - u|)$, for some given function J. It is elementary that there exists no critical phenomenon if $J(n) = 0$ for all large n. The picture is much more interesting if the interaction has infinite range. For the case when $J(n) \sim n^{-\alpha}$ as $n \to \infty$, it turns out that the process is totally ordered at all temperatures if $\alpha \le 1$, and disordered at all temperatures if $\alpha > 2$. There is a phase transition if $1 < \alpha \le 2$, in the sense that there exists a critical temperature above which there is no magnetization, and below which the spontaneous magnetization is non-zero (see Ruelle (1968), Dyson (1969a, b, 1971), and Fröhlich and Spencer (1982)). Furthermore, the phase transition is of the first order when $J(n) \sim n^{-2}$, which is to say that the spontaneous

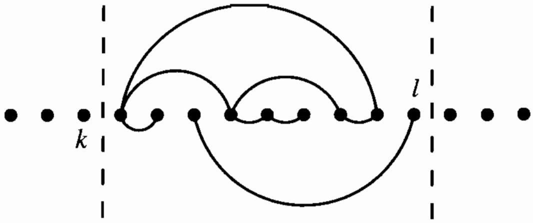


Figure 12.1. If A_k and A_l occur then no vertex in $\{k+1, k+2, \dots, l\}$ is joined to any vertex outside this set.

magnetization is a discontinuous function of temperature at the critical point (see Aizenman, Chayes, Chayes, and Newman (1988)).

What are the corresponding results for long-range percolation in one dimension? We may think of the probability vector $\mathbf{p}$ as being a vector $(p(n) : n \geq 1)$ of numbers lying in the interval $[0, 1)$, so that each pair $\{u, v\}$ of points in $\mathbb{Z}$ is joined by an (open) edge with probability $p(|v - u|)$. The resulting graph G is a random graph with vertex set $\mathbb{Z}$, and the problem is to relate the existence (or not) of infinite components to the particular choice of probability vector $\mathbf{p}$.

It is not surprising that all the components of G are a.s. finite if the $p(n)$ are sufficiently small. We describe two sufficient conditions for G to contain no infinite component. Let A_k be the event that no vertex u $(\leq k)$ is joined to any vertex v $(\geq k+1)$. The sequence $(I_{A_k} : -\infty < k < \infty)$ of indicator functions is stationary with trivial tail σ-field, and the common mean value of these random variables is $P(A_0)$, where P denotes the appropriate probability measure. There exist exactly n edges of the form $\langle u, u+n \rangle$ where $u \leq 0$ and $u + n \geq 1$, and therefore

$$P(A_0) = \prod_{n=1}^{\infty} \{1 - p(n)\}^n.$$

Suppose now that $\sum_n np(n) < \infty$; then $p(n) > \frac{1}{2}$ for only finitely many n, whence $\sigma = \sup\{p(n) : n \geq 1\} < 1$. There exists $\alpha < \infty$ such that $1 - x \geq e^{-\alpha x}$ for $x \in [0, \sigma)$, and therefore

$$P(A_0) \geq \exp\left(-\alpha \sum_{n=1}^{\infty} np(n) \right) > 0.$$

We apply the ergodic theorem to the stationary sequence of indicator functions to deduce that infinitely many of the A_k occur, with probability 1, and thus all components of G are finite (see Figure 12.1). We have proved that G contains a.s. no infinite component if

$$\sum_{n=1}^{\infty} np(n) < \infty. \tag{12.1}$$

For the second such condition, we compare the component at the origin with a branching process; it is not difficult to make the following argument rigorous. The mean number of vertices which are adjacent to the origin is $\mu = 2\sum_n p(n)$. Each such vertex v is adjacent to a certain number of new vertices (that is, vertices which are not adjacent to the origin), and the mean number of such new vertices is at most μ. We may build the component at the origin by generations; in the resulting branching tree, each point has a family having mean size not exceeding μ. Therefore the size of the component at the origin is no larger in distribution than the size of a certain branching process with mean family-size μ. If $\mu \le 1$, such a branching process is (almost surely) finite, and the component at the origin follows suit; the special case when $\mu = 1$ poses no serious difficulty. It is a small step to deduce that all components of G are a.s. finite if

$$2\sum_{n=1}^{\infty} p(n) \le 1. \tag{12.2}$$

The above discussion appeared in Schulman (1983).

We turn next to conditions which are sufficient for G to contain an infinite component. The next result is not surprising when we remember that the mean degree of each point is $2\sum_n p(n)$.

(12.3) Theorem. *The graph G is almost surely connected if and only if*

(12.4) *the greatest common divisor of* $\{n : p(n) > 0\}$ *is* 1, *and*

$$\sum_{n=1}^{\infty} p(n) = \infty. \tag{12.5}$$

Condition (12.4) requires that the probability vector $\mathbf{p}$ be 'aperiodic' in the sense that, for all pairs u, v of vertices, there is strictly positive probability that u and v are joined by a path of G; it is clear that this condition is necessary for G to be (almost surely) connected. Similarly, the necessity of (12.5) is fairly obvious. To see this, we note that the probability that any given vertex is isolated equals

$$\prod_{n=1}^{\infty} \{1 - p(n)\}^2,$$

which is strictly positive if $\sum_n p(n) < \infty$; it follows that, under this condition, G contains (almost surely) infinitely many isolated vertices. The core of Theorem (12.3) is the assertion that, subject to the aperiodicity of $\mathbf{p}$, the graph G is almost surely connected if $\sum_n p(n) = \infty$. That is to say, if $\sum_n p(n) = \infty$ then not only does G contain an infinite component, but *every vertex belongs to this component*.

Theorem (12.3) was proved first by Grimmett, Keane, and Marstrand (1984) in the context of long-range models in d dimensions where $d \ge 1$. There is a considerably simpler proof due to Kalikow (see Kalikow and Weiss (1988)). Kalikow reached a somewhat stronger conclusion than that of the above theorem: his argument implies that, subject to (12.4) and (12.5), the random graph *G restricted to the half-line* $\{0, 1, 2, \dots\}$ is almost surely connected. Using Kalikow's argument as a starting point, Kesten (1992) has extended Theorem (12.3) to graphs on half-spaces, quarter-spaces, and so on.

Theorem (12.3) is a corollary also of the uniqueness theorem for the infinite open cluster of long-range percolation on $\mathbb{L}^d$ where $d \geq 1$; see Aizenman, Kesten, and Newman (1987a, b), and the beautiful proof of Burton and Keane (1989) described in Section 8.2. The line of reasoning is the following. If $\sum_n p(n) = \infty$ then, by the second Borel–Cantelli lemma, each vertex has (almost surely) infinite degree; thus each vertex belongs a.s. to some infinite component. However, there is (almost surely) a unique infinite component, and thus all vertices belong (almost surely) to the same component.

We return to long-range percolation in one dimension, and specifically to the twilight zone containing probability vectors $\mathbf{p}$ satisfying

$$\sum_{n=1}^{\infty} np(n) = \infty, \qquad \tfrac{1}{2} < \sum_{n=1}^{\infty} p(n) < \infty.$$

We specialize to the class of vectors for which

$$p(n) \sim \beta n^{-\alpha} \qquad \text{as } n \to \infty \tag{12.6}$$

for some positive constants α and β. We have from the foregoing remarks that (almost surely)

$$\begin{array}{ll} \text{for } \alpha \leq 1, & G \text{ is connected,} \\ \text{for } \alpha > 2, & G \text{ has no infinite component.} \end{array}$$

As in the case of the one-dimensional Ising model, there is a critical phenomenon if $1 < \alpha \leq 2$. In order to formulate the result correctly, we note that both short-range and long-range interactions are relevant to the occurrence of a critical phenomenon, and we parametrize the process accordingly in the following somewhat arbitrary manner. Let α and β be positive constants and let p satisfy $0 \leq p < 1$. We consider a probability vector $\mathbf{p}$ satisfying

$$p(1) = p, \quad p(n) \sim \beta n^{-\alpha} \qquad \text{as } n \to \infty, \tag{12.7}$$

so that p measures the strength of interaction between nearest neighbours, and α and β measure the strengths of interactions over long distances. We write $\theta(p, \alpha, \beta)$ for the probability that the origin belongs to an infinite component of G, and we are concerned with the case $1 < \alpha \leq 2$.

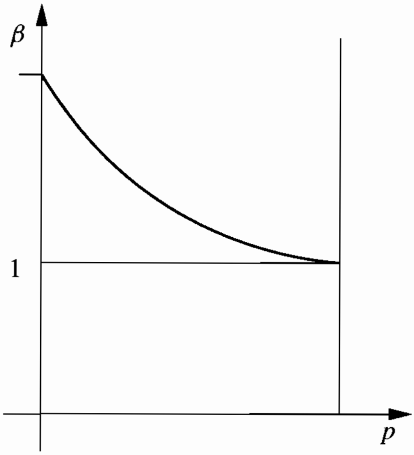


Figure 12.2. The curve $p = p_c(2, \beta)$. We have that $\theta(p, 2, \beta) = 0$ beneath the curve, and $\theta(p, 2, \beta) \geq \beta^{-1/2}$ above the curve. Thus θ is discontinuous as we cross the critical curve.

(12.8) Theorem. *Suppose that* $1 < \alpha < 2$ *and* $\beta > 0$. *There exists a critical value* $p_c(\alpha, \beta)$ *satisfying* $0 < p_c(\alpha, \beta) < 1$ *such that*

$$\theta(p, \alpha, \beta) \begin{cases} = 0 & \text{if } p < p_c(\alpha, \beta), \\ > 0 & \text{if } p > p_c(\alpha, \beta). \end{cases}$$

(12.9) Theorem. *Suppose that* $\alpha = 2$.

(a) *If* $\beta \leq 1$, *then* $\theta(p, 2, \beta) = 0$ *for all* $p < 1$.
(b) *If* $\beta > 1$, *there exists a critical value* $p_c(2, \beta)$ *satisfying* $0 < p_c(2, \beta) < 1$ *such that*

$$\theta(p, 2, \beta) \begin{cases} = 0 & \text{if } p < p_c(2, \beta), \\ \geq \beta^{-1/2} & \text{if } p > p_c(2, \beta). \end{cases}$$

Thus there is a critical phenomenon if either $1 < \alpha < 2$ or $\alpha = 2$, $\beta > 1$. In the latter case, the phenomenon is of the first order in the sense that $\theta(p, 2, \beta)$ is a discontinuous function of p at the critical point $p_c(2, \beta)$. These results are due to Newman and Schulman (1986) and Aizenman and Newman (1986), from the latter of which we reproduce in Figure 12.2 a picture of the phase diagram when $p(n) \sim \beta n^{-2}$.

One of the principal purposes of this section is to include Kalikow's contribution to the proof of Theorem (12.3).

Partial proof of Theorem (12.3). We restrict ourselves to proving that G is almost surely connected if

(12.10) $$p(1) > 0$$

and

(12.11) $$\sum_{n=1}^{\infty} p(n) = \infty.$$

Inequality (12.10) is the simplest condition which implies the aperiodicity of the probability vector **p**; we refer the reader to Grimmett, Keane, and Marstrand (1984) for the extra details in the general case. It is not hard to make rigorous the remarks after the statement of the theorem, and for that reason we omit also the proof that (12.11) is necessary for the connectedness of G. In partial compensation for these omissions we shall prove a stronger conclusion than is required, namely that (12.10) and (12.11) imply the a.s. connectedness of the random graph G^+ on the restricted vertex set $\{0, 1, 2, \dots\}$ with edges distributed at random in the usual way. The result for the doubly infinite vertex set $\mathbb{Z}$ will be an immediate consequence, by the following argument. If the two random graphs on the vertex sets $\{0, 1, 2, \dots\}$ and $\{\dots, -3, -2, -1\}$ are connected, then G is disconnected if and only if there exists no edge of the form $\langle u, v\rangle$ with $u < 0$ and $v \geq 0$. This event has probability

$$\prod_{n=1}^{\infty} \{1 - p(n)\}^n,$$

which equals 0, by assumption (12.11).

Let G_n be the graph G restricted to the vertex set $\{0, 1, 2, \dots, n\}$, and let A_n be the event that the vertices 0 and 1 belong to the same component of G_n; that is to say, $A_n = \{0 \leftrightarrow 1 \text{ in } G_n\}$. We note that $A_n \subseteq A_{n+1}$ for all n, so that

$$P(A_n) \to P(\lim A_n) = P(0 \leftrightarrow 1 \text{ in } G^+) \quad \text{as } n \to \infty.$$

We shall prove that

(12.12) $$P(A_n) \to 1 \qquad \text{as } n \to \infty,$$

implying that there exists almost surely a path from 0 to 1 in G^+; the required conclusion will follow quickly from this. We write B_n for the event that there exists in G_n some edge of the form $\langle 0, i\rangle$ for $2 \leq i \leq n$; the probability that B_n does not occur is

(12.13) $$P(\overline{B_n}) = \prod_{i=2}^{n} \{1 - p(i)\},$$

which satisfies

(12.14) $$P(\overline{B_n}) \to 0 \qquad \text{as } n \to \infty,$$

by (12.11).

Now,

$$P(\overline{A_n}) \le P(0 \nleftrightarrow 1 \text{ in } G'_n), \tag{12.15}$$

where G'_n is the random graph obtained from G_n by removing the edge $\langle 0, 1\rangle$ if it exists; we note that G'_n does not depend on the existence or non-existence of the edge $\langle 0, 1\rangle$. It is the case that

$$P(\overline{A_n}) \le P(0 \nleftrightarrow 1 \text{ in } G'_n,\ B_n) + P(\overline{B_n}), \tag{12.16}$$

of which the last term is controlled in (12.14). To control the first term, we argue as follows. If the event $\{0 \nleftrightarrow 1 \text{ in } G'_n,\ B_n\}$ occurs, and in addition the edge $\langle 0, 1\rangle$ exists, then G_n has strictly fewer components than the number C_{n-1} of components of the random graph on the smaller vertex set $\{1, 2, 3, \dots, n\}$. Writing K_n for the number of components of G_n, we have that

$$\begin{aligned} p(1)P(0 \nleftrightarrow 1 \text{ in } G'_n,\ B_n) &\le P(K_n < C_{n-1}) \\ &= P(K_n < K_{n-1}), \end{aligned}$$

since the pair (K_n, C_{n-1}) has the same joint distribution as the pair (K_n, K_{n-1}). We substitute this into (12.16), remembering that the K_n are integer-valued, to obtain

$$P(\overline{A_n}) \le \frac{1}{p(1)} P(K_{n-1} - K_n \ge 1) + P(\overline{B_n}). \tag{12.17}$$

Now $K_{n-1} - K_n \ge -1$, since the inclusion of one new vertex increases the number of components by 1 at most. Therefore,

$$\begin{aligned} P(K_{n-1} - K_n \ge 1) &= \sum_{i=1}^{\infty} P(K_{n-1} - K_n = i) \\ &\le E(K_{n-1} - K_n) + P(K_{n-1} - K_n = -1) \\ &\le E(K_{n-1} - K_n) + P(\overline{B_n}), \end{aligned} \tag{12.18}$$

since $K_{n-1} - K_n = -1$ if and only if vertex n is isolated in G_n, an event having probability $\{1 - p(1)\}P(\overline{B_n})$.

We wish to take the limit in (12.17) and (12.18) as $n \to \infty$. It is the case that

$$\liminf_{n\to\infty} E(K_{n-1} - K_n) \le 0 \tag{12.19}$$

since, if $\liminf_{n\to\infty} E(K_{n-1} - K_n) > \varepsilon$ where $\varepsilon > 0$, then $E(K_{n-1} - K_n) > \frac{1}{2}\varepsilon$ for all but finitely many values of n, which implies that

$$E(K_n) < E(K_N) - \tfrac{1}{2}\varepsilon(n - N)$$

for some N and all $n \geq N$; this is impossible since $E(K_n) \geq 0$ for all n. We choose a subsequence along which $E(K_{n-1}-K_n)$ approaches $\liminf_{n\to\infty} E(K_{n-1}-K_n)$, and we take the limit along this subsequence to obtain from (12.17) and (12.18) that

$$\begin{aligned}\liminf_{n\to\infty} P(\overline{A_n}) &\leq \frac{1}{p(1)} \liminf_{n\to\infty} E(K_{n-1} - K_n) \qquad \text{by (12.14)}\\ &\leq 0.\end{aligned}$$

However, $\lim_{n\to\infty} P(\overline{A_n})$ exists, since $\overline{A_{n+1}} \subseteq \overline{A_n}$ for all n, and thus (12.12) is proved.

Only a detail remains. If G^+ is disconnected, there exists a pair $x, x+1$ of non-negative integers such that x and $x+1$ are in different components of G^+. Such a pair $x, x+1$ has the property that they are not joined by a path in the subgraph of G^+ on the vertex set $\{x, x+1, x+2, \dots\}$, and therefore

$$\begin{aligned}P(G^+ \text{ is disconnected}) &\leq \sum_{x=0}^{\infty} P\big(x \nleftrightarrow x+1 \text{ in } \{x, x+1, x+2, \dots\}\big)\\ &= \sum_{x=0}^{\infty} P\big(0 \nleftrightarrow 1 \text{ in } \{0, 1, 2, \dots\}\big)\end{aligned}$$

by translation invariance. The final summand equals 0, and the proof of the claim is complete. □

12.4 Surfaces in Three Dimensions

We return to bond percolation on $\mathbb{L}^d$. The main weapon in two dimensions is planar duality. Using duality, we have seen that the critical point $p_c(2)$ marks the crossover between the existence of infinite closed clusters on the dual lattice and the existence of infinite open clusters on the primal lattice. The picture is not so clear in three dimensions.

Whereas in two dimensions every finite open cluster is surrounded by a closed *circuit* of the dual, such clusters in three dimensions are surrounded by closed *surfaces* on an appropriately defined lattice. We construct the dual lattice $\mathbb{L}^3_{\mathrm{d}}$ of $\mathbb{L}^3$ as follows. We take as vertices of $\mathbb{L}^3_{\mathrm{d}}$ the set

$$\mathbb{Z}^3 + (\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}) = \big\{(x_1 + \tfrac{1}{2}, x_2 + \tfrac{1}{2}, x_3 + \tfrac{1}{2}) : x \in \mathbb{Z}^3\big\}$$

of centres of the fundamental unit cubes of $\mathbb{L}^3$. We join two such vertices u and v by an edge whenever $\sum_{i=1}^{3} |u_i - v_i| = 1$. Thus $\mathbb{L}^3_{\mathrm{d}}$ is isomorphic to $\mathbb{L}^3$, and each edge of $\mathbb{L}^3$ passes through the centre of a unit face of $\mathbb{L}^3_{\mathrm{d}}$; following Aizenman,

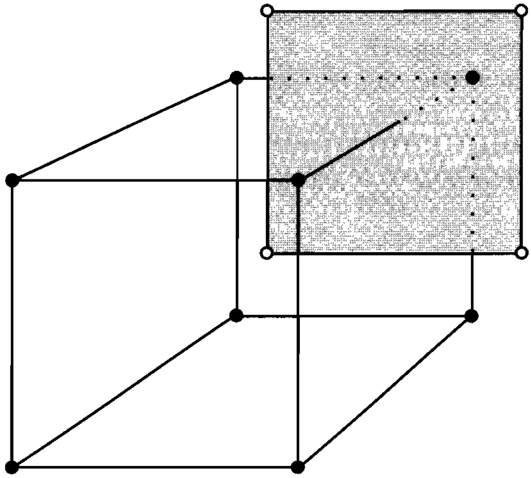


Figure 12.3. An elementary cube in $\mathbb{L}^3$ together with a dual plaquette. The open circles are vertices of the dual lattice, placed at the centres of the elementary cubes. Each edge of $\mathbb{L}^3$ passes through the centre of a face of the dual lattice.

Chayes, Chayes, Fröhlich, and Russo (1983), we call such faces *plaquettes*. Figure 12.3 is a sketch of a small part of $\mathbb{L}^3$ and its dual.

Let us consider bond percolation on $\mathbb{L}^3$ with edge-probability p. We declare each plaquette of the dual lattice $\mathbb{L}^3_\mathrm{d}$ to be *open* (respectively *closed*) if the intersecting edge of $\mathbb{L}^3$ is open (respective closed). Whereas in two dimensions we were interested in the existence of closed circuits in the dual, in three dimensions we study instead the existence of surfaces of closed plaquettes.

Each plaquette π has four bounding edges whose union we call the *boundary* of π, denoted by $\partial\pi$. Every collection $\Pi = (\pi_i : 1 \le i \le m)$ of distinct plaquettes determines a *surface*

$$S_\Pi = \bigcup_{i=1}^{m} \pi_i$$

together with a *boundary* ∂S_Π, defined as the set of edges of $\mathbb{L}^3_\mathrm{d}$ which belong to an odd number of plaquettes in Π. For the sake of brevity and concreteness, we shall study here the existence of surfaces of the following form only. Let c_n be the square boundary of the set $\{(x_1 + \frac{1}{2}, x_2 + \frac{1}{2}, \frac{1}{2}) : 0 \le x_1, x_2 \le n\}$ of vertices of $\mathbb{L}^3_\mathrm{d}$. We denote by C_n the event that there exists a collection Π of closed plaquettes whose boundary ∂S_Π is c_n. Thus C_n is the event that the circuit c_n is 'spanned' by a surface of closed plaquettes. Certain topological questions raise their heads at this point: do we want merely 'surfaces', or might we be more interested in orientable surfaces, or perhaps surfaces homeomorphic to the unit disc? We shall not discuss such questions here, but refer the reader to Aizenman, Chayes, Chayes, Fröhlich, and Russo (1983) and to Section 12.5.

How does the probability of C_n behave as a function of the parameter p? It is easy to see that

$$P_p(C_n) \geq (1-p)^{n^2}, \tag{12.20}$$

since C_n occurs whenever all the n^2 plaquettes in the smallest surface spanning c_n are closed. On the other hand,

$$P_p(C_n) \leq (1-p^4)^{4(n-1)}, \tag{12.21}$$

since, if C_n occurs, then at least one of the four plaquettes abutting each edge of c_n is closed. The bounds in (12.20) and (12.21) have qualitatively different asymptotic behaviours as $n \to \infty$, the lower bound decaying exponentially at a rate having the order of the *area* n^2 of c_n, and the upper bound at a rate having the order of its *perimeter length* $4n$. What is the true asymptotic behaviour of $P_p(C_n)$ for a given value of p? Closed surfaces of $\mathbb{L}^3_{\mathrm{d}}$ cut off open paths of $\mathbb{L}^3$, and we are led to conjecture that the answer to the question depends on the existence (or not) of an infinite open cluster in the primal lattice $\mathbb{L}^3$.

(12.22) Theorem.

(a) Perimeter law. *If* $0 < p < p_{\mathrm{c}}$, *there exists* $\alpha(p)$ *satisfying* $0 < \alpha(p) < \infty$ *such that*

$$P_p(C_n) \approx e^{-n\alpha(p)} \qquad \text{as } n \to \infty.$$

(b) Area law. *If* $p_{\mathrm{c}} < p < 1$, *there exists* $\beta(p)$ *satisfying* $0 < \beta(p) < \infty$ *such that*

$$P_p(C_n) \approx e^{-n^2\beta(p)} \qquad \text{as } n \to \infty.$$

In this theorem, p_{c} denote the critical probability of bond percolation on the lattice $\mathbb{L}^3$. We recall that '$a_n \approx b_n$' means that $\log a_n / \log b_n \to 1$.

Thus $\log P_p(C_n)$ is roughly proportional to the area of c_n when the density $1-p$ of closed plaquettes is small, and to the length of its perimeter when the density is large. The critical density of closed plaquettes is $1-p_{\mathrm{c}}$.

There is another percolation-type phenomenon associated with plaquettes: the formation of infinite connected clusters of closed plaquettes (with the obvious notion of adjacency). It is not difficult to see that this phenomenon occurs at a critical plaquette-density which is quite different from $1-p_{\mathrm{c}}$. By taking a horizontal slice through the dual lattice, we may satisfy ourselves that the critical density for the formation of infinite clusters of plaquettes is no more than $\frac{1}{2}$, whereas $p_{\mathrm{c}} < \frac{1}{2}$ and therefore $1 - p_{\mathrm{c}} > \frac{1}{2}$.

The results of this section appeared in Aizenman, Chayes, Chayes, Fröhlich, and Russo (1983), who considered also their generalizations to higher dimensions. They proved the area law for values of p exceeding the limit of the critical probabilities of slabs. See Kesten (1986a, 1987a) and Section 12.5 for related work.

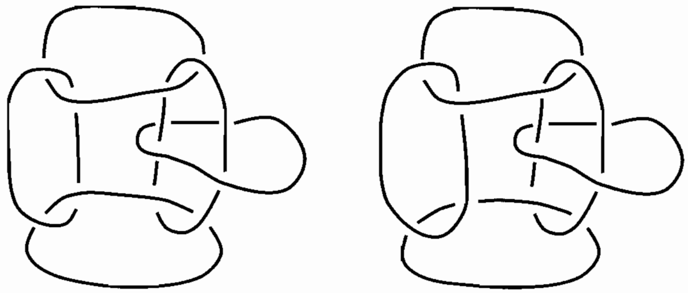


Figure 12.4. Sketches of two graphs. The first is entangled, the second is not.

12.5 Entanglement in Percolation

The emphasis so far has been upon the property of *connectedness*, and particularly upon whether or not there exists an infinite *connected* open set. There exist other geometrical properties of interest, including those of entanglement and rigidity. This section is devoted to 'entanglement in percolation'. For topological reasons which the reader is encouraged to consider, the following discussion is restricted to graphs embedded in the three-dimensional space $\mathbb{R}^3$.

Here is some terminology. Let $E \subseteq \mathbb{E}^3$. With each edge e ($\in E$) we associate the closed straight line segment of $\mathbb{R}^3$ joining the endpoints of e, and we denote by $[E]$ the union of all such line segments. We shall use the term 'sphere' to mean a subset of $\mathbb{R}^3$ which is homeomorphic to the unit sphere $\{x \in \mathbb{R}^3 : \rho(x) = 1\}$, where ρ denotes the Euclidean norm on $\mathbb{R}^3$. The complement of any sphere S has two connected components; we refer to the bounded component as the *inside* of S, denoted $\text{ins}(S)$, and the unbounded component as the *outside* of S, denoted $\text{out}(S)$.

There is a natural definition of the term 'entanglement' when applied to finite sets of edges of the lattice $\mathbb{L}^3$, namely the following. We call the finite edge-set E *entangled* if, for any sphere S not intersecting $[E]$, either $[E] \subseteq \text{ins}(S)$ or $[E] \subseteq \text{out}(S)$. Thus entanglement is a property of edge-sets rather than of graphs. However, with any edge-set E we may associate the graph $G(E)$ having edge-set E together with all incident vertices. Graphs $G(E)$ arising in this way have no isolated vertices. We call $G(E)$ *entangled* if E is entangled, and we note that $G(E)$ is entangled whenever it is connected. Figure 12.4 contains examples of entangled and unentangled graphs.

There are several possible ways of extending the notion of entanglement to infinite subgraphs of $\mathbb{L}^3$, and these ways are not equivalent. For the sake of being definite, we adopt here the following definition. Let E be an infinite subset of $\mathbb{E}^3$. We call E *entangled* if, for any finite subset F ($\subseteq E$), there exists a finite entangled subset E' of E such that $F \subseteq E'$. We call the infinite graph $G(E)$, defined as above, *entangled* if E is entangled, and we note that $G(E)$ is entangled whenever it is connected. The first graph in Figure 12.5 is entangled, but the second is not.

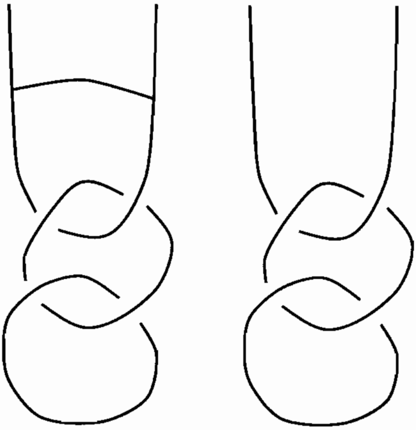


Figure 12.5. Examples of an infinite entangled graph and an infinite unentangled graph. The four lines at the top of the figure stretch off to infinity.

If we wish to include graphs such as the second within the definition, then the definition would need to be changed. It turns out that the results described in this section are valid for any reasonable such change, and we omit further discussion of this point.

Turning to percolation, we declare each edge of $\mathbb{L}^3$ to be open with probability p, in the usual way, and we call a subgraph G of $\mathbb{L}^3$ *open* if all its edges are open. We say that the origin 0 lies in an infinite open entanglement if there exists an infinite entangled set E of open edges at least one of which has 0 as an endvertex. We concentrate on the event

$$I = \big\{0 \text{ lies in an infinite open entanglement}\big\},$$

and the *entanglement probability*

$$\theta^{\text{ent}}(p) = P_p(I).$$

Since I is an increasing event, θ^{ent} is a non-decreasing function, whence

$$\theta^{\text{ent}}(p)\begin{cases} = 0 & \text{if } p < p_c^{\text{ent}}, \\ > 0 & \text{if } p > p_c^{\text{ent}}, \end{cases}$$

where the 'entanglement critical probability' p_c^{ent} is given by

$$p_c^{\text{ent}} = \sup\{p : \theta^{\text{ent}}(p) = 0\}. \tag{12.23}$$

Since every connected graph is entangled, it is immediate that $\theta(p) \le \theta^{\text{ent}}(p)$, whence $0 \le p_c^{\text{ent}} \le p_c$.

(12.24) Theorem. *The following strict inequalities are valid*:

(12.25) $$0 < p_c^{\text{ent}} < p_c.$$

Entanglements in percolation appear to have been studied first by Kantor and Hassold (1988). They used numerical techniques in order to obtain the estimate

$$p_c - p_c^{\text{ent}} \simeq 1.8 \times 10^{-7},$$

implying the conjecture that $p_c^{\text{ent}} < p_c$. This inequality was proved by Aizenman and Grimmett (1991) using the arguments laid down in Section 3.3.

Holroyd (1998b) has proved the strict positivity of p_c^{ent}. Whereas such inequalities are usually proved by straightforward counting arguments (cf. Theorem (1.10)), the required proof is complicated substantially by the difficulty of bounding the number of finite entangled graphs having n edges which contain the origin.

There remain several open problems concerning infinite entangled graphs, and we mention just two of these. First, prove that there exists almost surely a unique maximal infinite entangled graph whenever $\theta^{\text{ent}}(p) > 0$. Secondly, prove that the maximal entangled graph D containing the origin satisfies

$$P_p(|D| \geq n) \leq e^{-n\alpha(p)} \qquad \text{for all } n,$$

for some function α satisfying $\alpha(p) > 0$ when $p < p_c^{\text{ent}}$. The best result known to date has the general form

(12.26) $$P_p(|D| \geq n) \leq \exp\big(-\alpha(p)n/\log n\big),$$

where $\alpha(p) > 0$ for sufficiently small positive values of p. The logarithm in (12.26) may be replaced by any iterated logarithm, with an appropriate value of α. Such inequalities have been proved by Grimmett and Holroyd (1998), where further results and discussion concerning entanglements may be found. See also Grimmett (1999c).

12.6 Rigidity in Percolation

Attractive mathematical and physical problems arise in the study of rigid networks, and it is interesting to pose such questions in the context of percolation. Here are some fundamental definitions. Let $G = (V, E)$ be a finite graph. An *embedding* of G into $\mathbb{R}^d$ is an injection $f : V \to \mathbb{R}^d$; a *framework* (G, f) is a graph G together with an embedding f. A *motion* of a framework (G, f) is a differentiable family $\mathbf{f} = (f_t : 0 \leq t \leq 1)$ of embeddings of G, containing f, which preserves all edge lengths; that is, we require in particular that $f = f_T$ for some T, and that

(12.27) $$\rho\big(f_t(u) - f_t(v)\big) = \rho\big(f_0(u) - f_0(v)\big)$$

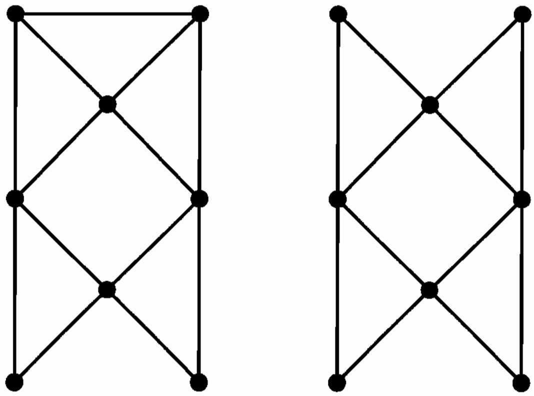


Figure 12.6. The first graph is rigid, the second is not.

for all edges $\langle u, v\rangle \in E$, where ρ denotes the Euclidean norm on $\mathbb{R}^d$. We call the motion $\mathbf{f}$ *rigid* if (12.27) holds for *all* pairs $u, v \in V$ rather than adjacent pairs only. A framework is called *rigid* if all its motions are rigid motions.

The above definition depends on the initial embedding f. However, for given dimension d, it turns out that the property of rigidity is in a sense 'generic' with respect to f. We do not explain this in detail but refer the reader to Gluck (1974), Graver, Servatius, and Servatius (1993), and Holroyd (1998a). In particular, it may be proved that there exists a natural probability measure μ on the set of embeddings of G such that: either (G, f) is rigid for μ-a.e. embedding f, or (G, f) is not rigid for μ-a.e. embedding. We call G *rigid* if the former holds.

We note that the rigidity of a graph depends on the number d of dimensions of the space in which it is to be embedded. In addition, rigid graphs are necessarily connected, but there exist connected graphs which are not rigid. See Figure 12.6 for two examples.

We turn now to the rigidity of infinite graphs. We call an infinite graph $G = (V, E)$ *rigid* if every finite subgraph of G is contained in some finite rigid subgraph of G.

Let $\mathcal{L}$ be a lattice in d dimensions, and consider bond percolation on $\mathcal{L}$ having edge-density p. [For a formal definition of a lattice, we may take that of either Grimmett (1978b) or Kesten (1982). See Section 12.1.] We call a subgraph of $\mathcal{L}$ *open* if all its edges are open. Let R be the event that the origin belongs to some infinite open rigid graph; we define the *rigidity probability* by

$$\theta^{\text{rig}}(p) = P_p(R).$$

Since R is an increasing event, θ^{rig} is a non-decreasing function, whence

$$\theta^{\text{rig}}(p) \begin{cases} = 0 & \text{if } p < p_c^{\text{rig}}, \\ > 0 & \text{if } p > p_c^{\text{rig}}, \end{cases}$$

where the 'rigidity critical probability' p_c^{rig} is given by

$$p_c^{rig} = p_c^{rig}(\mathcal{L}) = \sup\{p : \theta^{rig}(p) = 0\}.$$

Since rigid graphs are connected, we have that $\theta^{rig}(p) \leq \theta(p)$. This implies the inequality $p_c^{rig}(\mathcal{L}) \geq p_c(\mathcal{L})$, where $p_c(\mathcal{L})$ denotes the usual bond critical probability of $\mathcal{L}$.

(12.28) Theorem. *Let $\mathcal{L}$ be a d-dimensional lattice, where $d \geq 2$.*

(a) *We have the strict inequality $p_c(\mathcal{L}) < p_c^{rig}(\mathcal{L})$.*

(b) *It is the case that $p_c^{rig}(\mathcal{L}) < 1$ if and only if $\mathcal{L}$ is rigid.*

The methods required to prove this theorem may be found in Holroyd (1998a). Part (a) is proved using the techniques of Section 3.3, but utilizing a 'diminishment' rather than an 'enhancement'. It may be shown that the cubic lattice $\mathbb{L}^d$ is not rigid for any $d \geq 2$, implying by part (b) that $p_c^{rig}(\mathbb{L}^d) = 1$ for $d \geq 2$.

Combinatorial techniques permit a deeper study when $d = 2$. For the sake of definiteness, consider bond percolation on the triangular lattice $\mathbb{T}$, which may be shown to be a rigid lattice. A maximal open rigid subgraph of $\mathbb{T}$ is called an open rigid *component*. Holroyd (1998a) has proved that there exists almost surely a unique infinite open rigid component for any value of $p > p_c^{rig}(\mathbb{T})$ at which the function θ^{rig} is either left-continuous or right-continuous. This implies the uniqueness of the infinite rigid component for almost all p exceeding p_c^{rig}, and in addition that θ^{rig} is left-continuous wherever it is right-continuous. It is an open problem to prove continuity on the entire interval $(p_c^{rig}, 1]$, and there is doubt over whether or not to believe in the continuity of θ^{rig} *at* the critical point p_c^{rig}.

The study of the rigidity of percolation clusters was initiated by Jacobs and Thorpe (1995, 1996).

12.7 Invasion Percolation

Suppose that fluid is pumped under pressure into a random medium, and that each unit of the fluid displaces the least resistant unit of neighbouring background material. The following dynamic invasion process has received some attention as a model for such a situation. Let $(X(e) : e \in \mathbb{E}^d)$ be independent random variables indexed by the edge set of $\mathbb{L}^d$, each having the uniform distribution on $[0, 1]$. We construct a sequence $\mathbf{C} = (C_i : i \geq 0)$ of random connected subgraphs of the lattice in the following way. The graph C_0 contains the origin and no edges. Having defined C_i, we obtain C_{i+1} by adding to C_i the edge e_{i+1} for which $X(e_{i+1}) = \min\{X(e) : e \in \Delta C_i\}$; that is to say, we examine the edges in the edge boundary of C_i, and we add to C_i the edge f for which $X(f)$ is a minimum, together with any endvertex of this edge not already belonging to C_i.

A principal object of study is the empirical distribution function of the set $(X(e_j) : 1 \leq j \leq n)$, and particularly the limiting behaviour of this function as $n \to \infty$. Thus we define

$$Q_n(y) = \frac{1}{n} \sum_{j=1}^{n} I_{\{X(e_j) \leq y\}},$$

the proportion of edges e in C_n for which $X(e) \leq y$. The following intuitive argument is useful in understanding the behaviour of Q_n for large n. We call an edge e *p-open* if $X(e) < p$. Let π be a little larger than p_c, so that the lattice $\mathbb{L}^d$ contains (almost surely) an infinite π-open cluster. Once the invading fluid has reached some vertex in this cluster, it need never subsequently invade any edge f for which $X(f) > \pi$, since it is provided with an infinite supply of preferable edges. This motivates the guess that

(12.29) $$Q_n(y) \to 1 \qquad \text{if } y > p_c.$$

On the other hand, as the fluid devours this infinite π-open cluster, the values of the edges therein have a conditional distribution which is uniform on $[0, \pi]$, suggesting that

(12.30) $$Q_n(y) \to y/p_c \qquad \text{if } y < p_c.$$

Chayes, Chayes, and Newman (1985) have proved these results, subject to one of the standard conjectures of percolation of their time. Taken in conjunction with Theorem (7.2), their results imply (12.29) and (12.30).

In practice, background material cannot be forced out of an edge by the invading fluid if the edge in question is already 'trapped' in a cavity of the invading cluster. 'Invasion percolation with trapping' takes this fact into account; see Chayes, Chayes, and Newman (1987).

For motivation and information see de Gennes and Guyon (1978), Lenormand and Bories (1980), Chandler, Koplick, Lerman, and Willemsen (1982), Wilkinson and Willemsen (1983), Chayes, Chayes, and Newman (1985, 1987), Kesten (1987e), and Pokorny, Newman, and Meiron (1990).

12.8 Oriented Percolation

Modellers of galaxies, semiconductors, and elementary particles have had reason to study percolation on oriented lattices, and in particular on the 'north–east' lattice $\vec{\mathbb{L}}^d$ obtained by orienting each edge of $\mathbb{L}^d$ in the direction of increasing coordinate-value (see Figure 12.7 for a two-dimensional illustration). There are many parallels between results for oriented percolation and those for ordinary percolation; on the other hand the corresponding proofs often differ greatly, largely

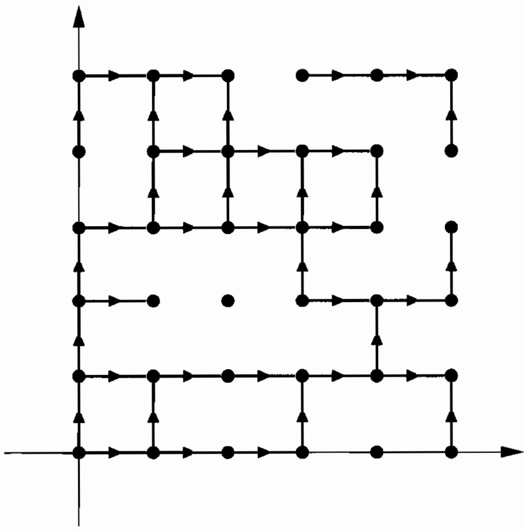


Figure 12.7. Part of the 'north–east' lattice in which each edge has been deleted with probability $1 - p$, independently of all other edges.

because the existence of one-way streets restricts the degree of spatial freedom of the traffic.

Suppose that $0 \leq p \leq 1$. We examine each edge of $\vec{\mathbb{L}}^d$ in turn, declaring it to be *open* with probability p and *closed* otherwise; as usual, the edge-states are independent of one another. We now supply fluid at the origin, but allow it to travel along open edges in the directions of their orientations only. We denote by C the set of vertices which may be reached from the origin along open directed paths; C contains exactly those vertices which are wetted by fluid supplied at the origin. The *percolation probability* is given by

$$\vec{\theta}(p) = P_p(|C| = \infty), \tag{12.31}$$

and the critical probability $\vec{p}_c(d)$ by

$$\vec{p}_c(d) = \sup\{p : \vec{\theta}(p) = 0\}. \tag{12.32}$$

It is an interesting exercise to show that $0 < \vec{p}_c(d) < 1$ if $d \geq 2$. The exact numerical value of $\vec{p}_c(d)$ is of course unknown; as for inequalities, we know that $0.6298 \leq \vec{p}_c(2) \leq \frac{2}{3}$; see Durrett (1984) and Liggett (1995b).

The process is understood quite well when $d = 2$, for which case Durrett (1984) has written a rather comprehensive survey. By looking at the set A_n of wet vertices on the diagonal $\{x \in \mathbb{Z}^2 : x_1 + x_2 = n\}$ of $\vec{\mathbb{L}}^2$, one may reformulate two-dimensional oriented percolation as a one-dimensional contact process in discrete time (see Liggett (1985, Chapter 6)). It turns out that $\vec{p}_c(2)$ may be characterized in terms of the velocity of the rightwards edge of a contact process on $\mathbb{Z}$ whose initial distribution places infectives to the left of the origin and susceptibles to the right. With the support of arguments from branching processes and ordinary percolation, one may prove such results as the exponential decay of the cluster size distribution when $p < \vec{p}_c(2)$, and its sub-exponential decay when $p > \vec{p}_c(2)$: there exist $\alpha(p), \beta(p) > 0$ such that

$$(12.33) \qquad e^{-\alpha(p)\sqrt{n}} \le P_p\big(n \le |C| < \infty\big) \le e^{-\beta(p)\sqrt{n}} \qquad \text{if } \vec{p}_c(2) < p < 1.$$

Progress in higher dimensions has been more limited. Cox and Durrett (1983) have explored the asymptotic behaviour of $\vec{p}_c(d)$ as $d \to \infty$, showing in particular that $\vec{p}_c(d) \sim d^{-1}$. More recent work is contained in Wierman (1985a), and in Durrett and Schonmann (1987, 1988a).

There is a close relationship between oriented percolation and the contact model (see Section 13.5), and methods developed for the latter model may often be applied to the former. In particular, the arguments of Bezuidenhout and Grimmett (1990, 1991) may be used to obtain several results including the fact that $\vec{\theta}(\vec{p}_c) = 0$ for general $d \ge 2$.

We close this section with an open problem of a different sort. Suppose that each edge of $\mathbb{L}^2$ is oriented in a random direction, horizontal edges being oriented eastwards with probability p and westwards otherwise, and vertical edges being oriented northwards with probability p and southwards otherwise. Let $\eta(p)$ be the probability that there exists an infinite oriented path from the origin. It is not hard to use Theorem (11.12) to show that $\eta(\frac{1}{2}) = 0$. We ask whether or not $\eta(p) > 0$ if $p \ne \frac{1}{2}$.

12.9 First-Passage Percolation

First-passage percolation is the half-brother of ordinary percolation. It was formulated by Hammersley and Welsh (1965) as a time-dependent model for the passage of fluid through a porous medium, and has since been a source of probabilistic problems endowed with both excellent motivation and mathematical appeal; some of the best of these questions remain partly unanswered. In addition, its birth witnessed the recognition of subadditive stochastic processes, of which the ergodic theorem is now a major weapon of standard issue to graduate students in probability. Only inadequate justice can be done to the subject in a few pages only, and we refer the reader to the monograph of Smythe and Wierman (1978) and the work of Kesten (1986a, 1987a, e) and Newman (1995) for more detailed information and historical references.

To each edge e of $\mathbb{L}^d$ we allocate a random *time coordinate* $T(e)$, which we think of as being the time required for fluid to flow along e; we assume that the $T(e)$ are independent non-negative random variables having common distribution function F. For any path π we define the *passage time* $T(\pi)$ of π by

$$T(\pi) = \sum_{e \in \pi} T(e).$$

The *first-passage time* $a(x, y)$ between vertices x and y is given by

$$a(x, y) = \inf\{T(\pi) : \pi \text{ a path from } x \text{ to } y\};$$

thus $a(x, y)$ is the time taken by fluid to wet y, having been supplied at x. If we supply an unlimited quantity of fluid at the origin, the set of wet vertices at time t is

$$W(t) = \{x \in \mathbb{Z}^d : a(0, x) \leq t\},$$

and the main questions pertain to the rate of growth of $W(t)$ as $t \to \infty$. The easiest way of describing the asymptotic behaviour of $W(t)$ involves re-scaling the lattice as $t \to \infty$, and the result is easiest stated for the set $\widetilde{W}(t)$ obtained from $W(t)$ by filling in the holes: we define

$$\widetilde{W}(t) = \{y + B : y \in W(t)\},$$

where $B = \left[-\frac{1}{2}, \frac{1}{2}\right]^d$ is the unit cube centred at the origin.

It turns out that $\widetilde{W}(t)$ grows approximately linearly as time passes, in the sense that, subject to a suitable moment condition on the time coordinate distribution, there exists a non-random limit set L for $t^{-1}\widetilde{W}(t)$ having non-empty interior such that either

(a) L is compact with

$$(1 - \varepsilon)L \subseteq \frac{1}{t}\widetilde{W}(t) \subseteq (1 + \varepsilon)L \qquad \text{eventually, almost surely,}$$

for all $\varepsilon > 0$, or

(b) $L = \mathbb{R}^d$ with

$$\frac{1}{t}\widetilde{W}(t) \supseteq \{x \in \mathbb{R}^d : |x| \leq M\} \qquad \text{eventually, almost surely,}$$

for all $M > 0$.

The earliest such 'shape theorem' was proved by Richardson (1973), and the last words were due to Cox and Durrett (1981) and Kesten (1986a). Few non-trivial facts are known about the shape of the limit set L; worthy of note in this regard is the finding of Durrett and Liggett (1981) that, for a closely related problem, the boundary of L may contain straight line segments when $d = 2$.

Specializing to the spread of fluid in an axial direction, we define

$$a_{mn} = a(e_m, e_n) \qquad \text{for } -\infty < m \leq n < \infty,$$

the first-passage time from $e_m = (m, 0, 0, \dots, 0)$ to $e_n = (n, 0, 0, \dots, 0)$. Two of the most useful properties of the family $\mathbf{a} = (a_{mn} : m \le n)$ are *stationarity*, in that $(a_{mn} : m \le n)$ and $(a_{m+1,n+1} : m \le n)$ have the same finite-dimensional distributions, and *subadditivity*, in that

$$a_{mn} \le a_{mr} + a_{rn} \qquad \text{whenever } m \le r \le n.$$

This subadditive inequality is a consequence of the fact that a_{mn} is no greater than the infimum of the passage times of paths from e_m to e_n which pass through e_r. Subject to a moment condition, we have by the subadditive ergodic theorem that

$$\frac{1}{n} a_{0n} \to \mu \qquad \text{almost surely}$$

for some constant $\mu = \mu(F, d)$ called the *time constant* of the process. This result is one of the main steps towards establishing the 'shape' theorem above. Little is known about the way in which $\mu(F, d)$ depends on the choice of the distribution function F.

The sequence $\mathbf{a}$ satisfies a large-deviation theorem (Grimmett and Kesten (1984a)), but it is an open problem to find sequences $(\alpha(n), \beta(n) : n \ge 0)$ such that $(a_{0n} - \alpha(n))/\beta(n)$ has a non-trivial limit distribution as $n \to \infty$. See Kesten (1986a, 1987e), and also Kesten and Zhang (1997) for a central limit theorem for a certain critical two-dimensional example.

Of major interest in recent years has been the study of fluctuations of the set $\widetilde{W}(t)$ for large t. See Alexander (1993b, 1997), Kesten (1993), Licea and Newman (1996), Licea, Newman, and Piza (1996), Newman (1995), and Newman and Piza (1995).

12.10 Continuum Percolation

Young lilies inhabit a beautiful large square pond. They grow at a uniform rate. How long need a snail wait beside the pond before it becomes possible for him to traverse the pond without getting his feet wet? This question may be rephrased in terms of the following fundamental problem of stochastic geometry: ascertain the minimal density of unit discs in the plane which guarantees the a.s. existence of an infinite cluster. It leads to a percolation-type process which has been dubbed by others the 'Poisson blob model' and the 'Boolean model'. A picture of the snail's route appears in Figure 12.8.

Let $(X_i : i \in I)$ be the points of a Poisson process with intensity λ in $\mathbb{R}^d$, where $d \ge 2$. For each $i \in I$, we position a closed d-dimensional sphere of radius ξ with its centre at X_i. Writing S_i for the sphere with centre X_i, we have that $S_i = S + X_i$, where $S = \{x \in \mathbb{R}^d : \rho(x) \le \xi\}$ and ρ is the Euclidean norm. We are interested in the shapes of clusters of overlapping spheres, and particularly in the possible existence of infinite clusters.

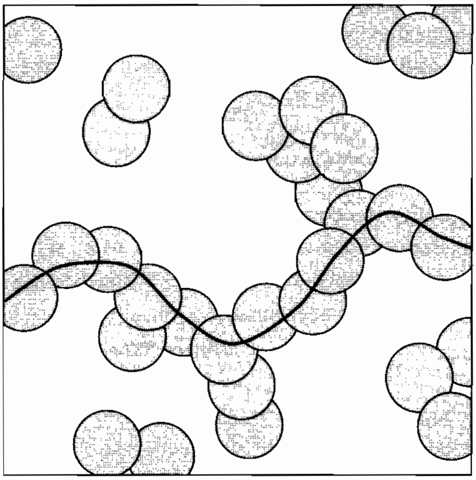


Figure 12.8. The thick line indicates a possible dry crossing of the lily pond. The centres of the circles are the points of a Poisson process. When the circles are sufficiently large, there exists a crossing of the box which is contained in the union of the circles.

We call two spheres S_i and S_j *adjacent* if $S_i \cap S_j \neq \varnothing$. We write $S_i \leftrightarrow S_j$ if there exists a sequence $S_{i_1}, S_{i_2}, \dots, S_{i_k}$ of spheres such that $S_{i_1} = S_i$, $S_{i_k} = S_j$, and S_{i_l} is adjacent to $S_{i_{l+1}}$ for $1 \leq l < k$. A *cluster* of spheres is a set $(S_i : i \in J)$ of spheres which is maximal with the property that $S_i \leftrightarrow S_j$ for all $i, j \in J$. We identify a cluster with the set of centres of the spheres in the cluster. The *size* of a cluster is thus the number of spheres belonging to it.

We wish to study the sizes and shapes of clusters for specified values of λ and ξ. From the re-scaling property of the Poisson process, the cluster shapes have the same distributions (except for a change in the length scale) whenever the parameters have the form $\alpha^{-d}\lambda$ and $\alpha\xi$ respectively, as α ranges over $(0, \infty)$. We may therefore assume without loss of generality that $\xi = 1$, so that the unique parameter of the process is the intensity λ of the point process.

From our experience of bond percolation on $\mathbb{L}^d$, we can be fairly sure that there exists a non-trivial critical intensity λ_c such that (almost surely) there exists an infinite cluster when $\lambda > \lambda_c$ and not when $\lambda < \lambda_c$. We shall prove this and rather more, although we do not attempt a complete description of the process here.

As for notation, we write P^λ for the probability measure associated with a Poisson process with intensity λ, and E^λ for the corresponding expectation operator. We denote by $W(x)$ the cluster containing the point x of $\mathbb{R}^d$; more formally, $W(x)$ is defined to be the set of points in the Poisson process which are centres of spheres in the cluster containing x. The *percolation probability* $\gamma(\lambda)$ is defined by

$$\gamma(\lambda) = P^{\lambda}(|W| = \infty), \tag{12.34}$$

where $W = W(0)$ is the cluster at the origin. It is clear from the argument leading to Theorem (1.11) that

$$P^{\lambda}\big(\text{there exists } x \text{ with } |W(x)| = \infty\big) = \begin{cases} 0 & \text{if } \gamma(\lambda) = 0, \\ 1 & \text{if } \gamma(\lambda) > 0. \end{cases}$$

(12.35) Theorem. *There exists λ_c satisfying $0 < \lambda_c < \infty$ such that*

$$\gamma(\lambda) \begin{cases} = 0 & \text{if } \gamma < \lambda_c, \\ > 0 & \text{if } \gamma > \lambda_c. \end{cases} \tag{12.36}$$

Furthermore, the mean cluster size $E^{\lambda}|W|$ is finite whenever $\lambda < \lambda_c$.

The proof of this theorem is due to Zuev and Sidorenko (1985a, b); see also Menshikov, Molchanov, and Sidorenko (1986) and Menshikov and Sidorenko (1987). The main idea is to approximate to the 'continuous' problem by site percolation problems on special lattices constructed by partitioning $\mathbb{R}^d$ into small cubes. This approximation technique may be used to explore many other properties of the continuous problem. For example, it is not difficult to show that

$$\zeta(\lambda) = \lim_{n \to \infty} \left\{ -\frac{1}{n} \log P^{\lambda}(|W| = n) \right\}$$

exists and is strictly positive when $0 < \lambda < \lambda_c$. Furthermore, there exists almost surely a *unique* infinite cluster of overlapping spheres whenever $\gamma(\lambda) > 0$.

Numerous generalizations of the problem spring readily to mind, and some are discussed by Zuev and Sidorenko (1985a, b) and Menshikov and Sidorenko (1987). Interesting phenomena may occur if the random sets are permitted to be large, and Hall (1985) has shown that there exist similar processes for which the mean cluster size and the percolation probability have quite different critical behaviours. Suppose that a sphere of random radius is positioned at each point of a Poisson process with intensity λ; we assume that the radii of the spheres are independent and identically distributed random variables. Hall has shown that there exists a non-trivial critical intensity λ_T for the divergence of the mean cluster size if and only if a typical sphere-radius R satisfies $E(R^{2d}) < \infty$; if $E(R^{2d}) = \infty$ then the mean cluster size is infinite for all $\lambda > 0$. On the other hand, if $E(R^{2d-1}) < \infty$ then all clusters are almost surely finite for all positive but sufficiently small values of λ. If $E(R^{2d-1}) < \infty$ but $E(R^{2d}) = \infty$, then the mean cluster size is infinite for all $\lambda > 0$ but there exists $\lambda_c > 0$ such that all clusters are almost surely finite when $\lambda < \lambda_c$.

The principal reference for continuum percolation is the systematic account of Meester and Roy (1996). Further references are given by Kesten (1987e), who discusses other applications to polymerization, cluster analysis, and communication systems. See Hall (1988) for an account of the general theory of 'coverage' processes.

Proof of Theorem (12.35). There is a natural way of relating the continuous-space process to an ordinary site percolation process. Let n be a positive integer and let $\mathbb{Z}_n^d$ be the set $n^{-1}\mathbb{Z}^d$; that is to say

$$\mathbb{Z}_n^d = \{(x_1/n, x_2/n, \ldots, x_d/n) : x \in \mathbb{Z}^d\}.$$

Later, we shall take the limit as $n \to \infty$. We partition $\mathbb{R}^d$ into cubes whose centres are at the points of $\mathbb{Z}_n^d$, defining

$$B_n(x) = \prod_{i=1}^{d} \left[x_i - \frac{1}{2n}, x_i + \frac{1}{2n}\right) \qquad \text{for } x \in \mathbb{Z}_n^d.$$

We turn $\mathbb{Z}_n^d$ into a lattice $\mathcal{L}_n$ by defining the adjacency relation $\sim$ on $\mathbb{Z}_n^d$ with the rule that $x \sim y$ if and only if there exist points $u \in B_n(x)$ and $v \in B_n(y)$ such that

$$\left(\sum_{i=1}^{d}(u_i - v_i)^2\right)^{1/2} \leq 2. \tag{12.37}$$

We shall consider site percolation on the ensuing lattice $\mathcal{L}_n$. We declare a vertex x of $\mathcal{L}_n$ to be *open* if there exist one or more points of the Poisson process within the cube $B_n(x)$, and *closed* otherwise. It is elementary that the states of different vertices are independent random variables, and that the probability $p_n(\lambda)$ that any given vertex is open is given by

$$\begin{aligned} p_n(\lambda) &= 1 - P_\lambda\big(B_n(0) \text{ contains no point}\big) \\ &= 1 - e^{-\lambda n^{-d}}. \end{aligned} \tag{12.38}$$

We shall make liberal use of the properties of site percolation on $\mathcal{L}_n$ with no greater proof than the imprecation to adapt the arguments given earlier for bond percolation on $\mathbb{L}^d$. Thus we shall assume for example that the mean cluster size is finite throughout the subcritical phase.

Each point X of the Poisson process lies in some unique elementary cube $B_n(x)$ of $\mathcal{L}_n$ and we write $x = \pi_n(X)$. The function π_n is a many–one mapping from the points of the Poisson process into the set $\mathbb{Z}_n^d$ of vertices of $\mathcal{L}_n$. If S_1 and S_2 are two intersecting spheres with centres X_1 and X_2, then the Euclidean distance between X_1 and X_2 is no greater than 2, giving by the definition of the adjacency relation of $\mathcal{L}_n$ that $\pi_n(X_1) \sim \pi_n(X_2)$; see Figure 12.9. It follows that, if the origin lies in

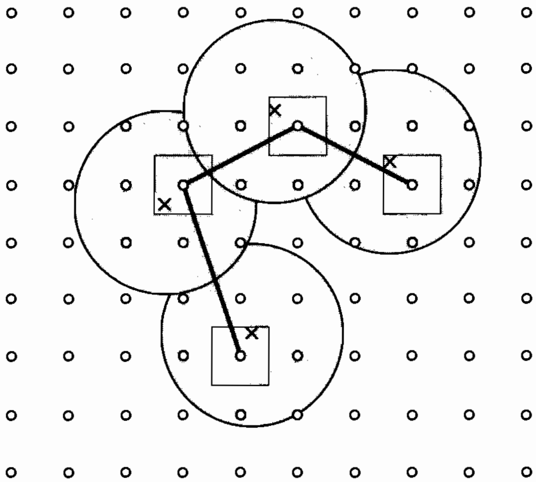


Figure 12.9. The crosses mark the centres of four overlapping circles. Each cross is associated with the vertex closest to it, and the adjacency relation of the lattice $\mathcal{L}_n$ has been chosen so that two such vertices are adjacent whenever the corresponding circles overlap.

an infinite cluster of spheres, then there exists an infinite open cluster in $\mathcal{L}_n$. In terms of the critical intensity

(12.39) $$\lambda_c = \sup\{\lambda : \gamma(\lambda) = 0\}$$

and the critical probability $p_c(\mathcal{L}_n)$ of site percolation on $\mathcal{L}_n$, we have proved that $p_n(\lambda) \geq p_c(\mathcal{L}_n)$ whenever $\lambda > \lambda_c$. We rewrite this implication with the aid of (12.38) as

(12.40) $$\text{`}\lambda > \lambda_c\text{'} \quad \text{implies} \quad \text{`}\lambda \geq -n^d \log\{1 - p_c(\mathcal{L}_n)\}\text{'},$$

which is to say that

(12.41) $$\lambda_c \geq -n^d \log\{1 - p_c(\mathcal{L}_n)\}.$$

However, $0 < p_c(\mathcal{L}_n) < 1$, and therefore $\lambda_c > 0$. Actually, the foregoing argument was largely for practice only, since inequality (12.41) is not itself of great value in studying the mean cluster size $E^\lambda|W|$. Of greater value would be the corresponding inequality for

(12.42) $$\lambda_T = \sup\{\lambda : E^\lambda|W| < \infty\}.$$

Clearly,

(12.43) $$\lambda_T \leq \lambda_c$$

and our principal object is to prove that equality holds here. In order to obtain the inequality for λ_{T} corresponding to (12.41), we need to relate the mean number of spheres in W to the mean size of an open cluster of $\mathcal{L}_n$. Let D_n be the set of open vertices of $\mathcal{L}_n$ which are joined by open paths to either the origin or some vertex in the set $\{x \in \mathbb{Z}_n^d : x \sim 0\}$ of neighbours of the origin, and suppose that $|D_n| = m$. Each vertex x of D_n is the centre of a cube $B_n(x)$ which contains at least one point of the Poisson process. The mean number of such points in any such cube is therefore $E(N \mid N \geq 1)$, where N is a random variable having the Poisson distribution with parameter λn^{-d}; hence

$$\begin{aligned} E^\lambda|\pi_n^{-1}(D_n)| &= \sum_{m=0}^{\infty} mE(N \mid N \geq 1)P^\lambda(|D_n| = m) \\ &= \frac{\lambda n^{-d}}{1 - e^{-\lambda n^{-d}}} E^\lambda|D_n| \end{aligned}$$

by an elementary calculation. The origin of $\mathcal{L}_n$ has at most $(4n)^d$ neighbours, and therefore

$$E^\lambda|D_n| \leq \big\{1 + (4n)^d\big\} E^\lambda|C_n|,$$

where C_n is the open cluster of $\mathcal{L}_n$ at the origin. It is the case that $W \subseteq \pi_n^{-1}(D_n)$ by the remarks prior to (12.39), and therefore

$$E^\lambda|W| \leq \frac{\lambda n^{-d}\{1 + (4n)^d\}}{1 - e^{-\lambda n^{-d}}} E^\lambda|C_n|.$$

It follows that $E^\lambda|C_n| = \infty$ whenever $E^\lambda|W| = \infty$, which is to say that

(12.44) $$\lambda_{\mathrm{T}} \geq -n^d \log\big\{1 - p_{\mathrm{c}}(\mathcal{L}_n)\big\}$$

as in (12.41); we have used the fact that $E^\lambda|C_n| = \infty$ only when $p \geq p_{\mathrm{c}}(\mathcal{L}_n)$.

We seek next an upper bound for λ_{c}, and intend to prove that

(12.45) $$\lambda_{\mathrm{c}} \leq -l_n^d n^d \log\big\{1 - p_{\mathrm{c}}(\mathcal{L}_n)\big\},$$

where $l_n = 1 + n^{-1}\sqrt{d}$. It is an immediate consequence of (12.43)–(12.45) that

(12.46) $$\lambda_{\mathrm{T}} = \lambda_{\mathrm{c}} = \lim_{n \to \infty} \big[-n^d \log\big\{1 - p_{\mathrm{c}}(\mathcal{L}_n)\big\}\big],$$

using the fact that $l_n \to 1$ as $n \to \infty$. Incidentally, (12.43)–(12.45) imply that $0 < \lambda_{\mathrm{c}} < \infty$.

We now prove (12.45). We have from the definition of the adjacency relation of $\mathcal{L}_n$ that

$$\left(\sum_{i=1}^{d} (u_i - v_i)^2\right)^{1/2} \leq 2l_n$$

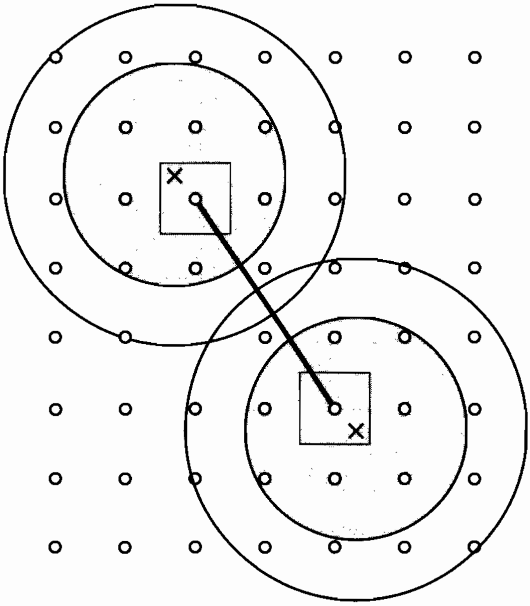


Figure 12.10. Here are two vertices of the lattice $\mathcal{L}_n$ which are adjacent even though there exist corresponding points (marked with crosses) which are separated by a distance exceeding 2. Unit circles centred at these points do not overlap, but concentric circles with increased radius $l_n = 1 + n^{-1}\sqrt{d}$ do indeed overlap.

whenever $u \in B_n(x)$, $v \in B_n(y)$, and $x \sim y$; just note that the diagonal of the typical elementary cube $B_n(0)$ has length $n^{-1}\sqrt{d}$. Suppose that we replace each sphere of radius 1 by a sphere of radius l_n. If x and y are adjacent open vertices of $\mathcal{L}_n$, then *all* such enlarged spheres with centres in $B_n(x)$ intersect *all* such enlarged spheres with centres in $B_n(y)$; see Figure 12.10. Therefore, the cluster of spheres (with radius l_n) at the origin is infinite whenever the open cluster of $\mathcal{L}_n$ at the origin is infinite. However, the process of spheres with radius l_n is equivalent to a process of unit-radius spheres with an amended intensity λl_n^d. We now see as in (12.40) that

$$`p_n(\lambda) > p_c(\mathcal{L}_n)' \quad \text{implies} \quad `\lambda l_n^d \geq \lambda_c',$$

which is to say that

$$-n^d \log\{1 - p_c(\mathcal{L}_n)\} \geq l_n^{-d}\lambda_c$$

as required in (12.45). □

Chapter 13

Percolative Systems

13.1 Capacitated Networks

The emphasis of first-passage percolation is upon flow rate subject to constraints of time. An alternative model involves constraints of capacity, and the classical model of this form is the 'capacitated network' of Ford and Fulkerson (1962). Suppose that we are provided with a finite graph G, of which each edge e has a non-negative *capacity* $c(e)$; we may think of $c(e)$ as an upper bound for the quantity of fluid which may flow along e in unit time. Let S and T be disjoint sets of vertices, called *source* and *sink* vertices respectively. A *flow* from S to T is an assignment of a non-negative number $f(e)$ and an orientation to each edge e of G such that

$$I(v) = \sum_{w:v\rightarrow w} f(\langle v, w\rangle) - \sum_{w:w\rightarrow v} f(\langle v, w\rangle) \tag{13.1}$$

satisfies $I(v) = 0$ for all vertices $v \notin S \cup T$, where the first (respectively second) summation is calculated over all neighbours w of v for which $\langle v, w\rangle$ is oriented away from (respectively towards) v; thus fluid is conserved at all vertices except possibly at sources and sinks. A flow is *admissible* if

$$f(e) \leq c(e) \qquad \text{for all edges } e,$$

and the *value* of such a flow is defined to be $\sum_{v\in S} I(v)$, the aggregate amount of fluid entering G at the source vertices. The *maximum flow* is the largest value of all admissible flows.

The max-flow min-cut theorem characterizes the maximal flow through the network in terms of the sizes of its cutsets. A set E of edges is called a (S, T)-*cutset* if all paths from S to T use at least one edge of E. The *size* of a (S, T)-cutset E is defined to be the sum of the capacities of the edges in E. It is clear that the

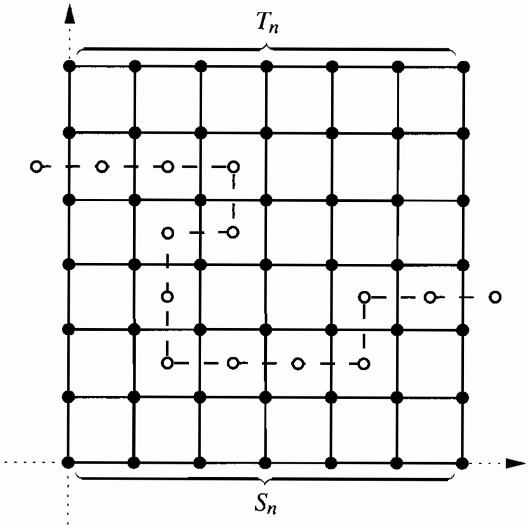


Figure 13.1. Every left–right crossing in the dual traverses edges in U_n which form a cutset. It is not difficult to see that every minimal cutset arises in such a way.

maximum flow cannot exceed the size of any cutset, and the max-flow min-cut theorem asserts that the maximum flow is actually equal to the minimum size of all (S, T)-cutsets.

We now specialize to the lattice $\mathbb{L}^d$. Suppose that the capacities $(c(e) : e \in \mathbb{E}^d)$ are independent random variables having the common distribution function F. We are interested in flows through large boxes, and denote by U_n the cube $[0, n]^d$ of $\mathbb{L}^d$. As sources we take the 'bottom' of U_n,

$$S_n = [0, n]^{d-1} \times \{0\},$$

and as sinks we take the 'top',

$$T_n = [0, n]^{d-1} \times \{n\}.$$

What can be said about the maximum flow φ_n from S_n to T_n through U_n? It is illustrative to consider in some detail the special case of two dimensions. By the max-flow min-cut theorem, φ_n equals the minimum size of (S_n, T_n)-cutsets. It is not difficult to see as in Figure 13.1 that each such cutset corresponds to a path in the dual graph which traverses the rectangle U_n from left to right. If to each edge of the dual lattice we allocate a time coordinate equal in value to the capacity of the edge which it crosses, then the size of such a cutset equals the passage time of the corresponding dual path. That is to say, the maximum flow from bottom to top in U_n equals the minimum passage time from left to right in the dual. We now appeal to the theory of first-passage percolation to deduce (after some work, see

Grimmett and Kesten (1984a)) that, as $n \to \infty$,

$$\frac{1}{n}\varphi_n \to \mu(F, 2) \qquad \text{almost surely,}$$

where $\mu(F, 2)$ is the time constant of first-passage percolation on $\mathbb{L}^2$ having time coordinate distribution F.

The picture in higher dimensions is less clear, although Kesten (1985, 1986a, 1987a) has made considerable progress towards establishing results analogous to those of first-passage percolation. Consider, for example, the case $d = 3$. The asymptotic behaviour of φ_n depends qualitatively upon the size of the atom of F at zero, since if a large number of edges have capacity 0 then it is likely that there exists no path from S_n to T_n having non-zero capacity. The critical value of $F(0)$ is likely to be $1 - p_c(3)$, where $p_c(3)$ is the critical probability of bond percolation on $\mathbb{L}^3$. Certainly if $F(0) > 1 - p_c(3)$ then $\varphi_n = 0$ for all large n (almost surely). It remains an open problem to establish the asymptotic behaviour of φ_n when $F(0) < 1 - p_c(3)$. One expects that φ_n is proportional to the cross-sectional area n^2 of U_n, in that $n^{-2}\varphi_n$ converges to a positive limit as $n \to \infty$, but this has been established only for values of $F(0)$ not exceeding $\frac{1}{27}$ (see Kesten (1987a)).

Further discussions may be found in Grimmett and Kesten (1984a), Grimmett (1985c), and Kesten (1987e).

13.2 Random Electrical Networks

Percolation theory is one of the principal tools for analysing disordered media, and there are few better examples of its use than in the study of disordered electrical networks. It may not be too difficult to calculate the effective electrical resistance of a block of either material A or material B, but what is the effective resistance of a mixture of the two materials? The answer depends, of course, on the way in which the composite block is constructed from the materials in question. If the composition is disordered, then the most practicable assumption may be that each individual component is made of a randomly chosen material. It is not known how to calculate the effective resistance of such large blocks even if we are prepared to assume the maximum amount of independence between components.

We make the discussion more concrete as follows. Let $d \geq 2$. Instead of the capacities of the last section, we allocate to each edge e of $\mathbb{L}^d$ an electrical resistance $r(e)$ satisfying $0 \leq r(e) \leq \infty$, and we assume that the $r(e)$ are independent and identically distributed. The value $r(e) = 0$ corresponds to a superconductor, and $r(e) = \infty$ corresponds to an insulator. We denote by U_n the cube $[0, n]^d$ of the lattice $\mathbb{L}^d$. We turn U_n into an electrical network by applying a potential difference between the bottom and top (hyper)faces of U_n,

$$S_n = [0, n]^{d-1} \times \{0\}, \quad T_n = [0, n]^{d-1} \times \{n\};$$

see Figure 1.3 for a two-dimensional illustration. We write R_n for the effective resistance of the network, and we seek to understand how the asymptotic properties of R_n depend upon the underlying distribution of the edge-resistances. It is reasonable to guess that R_n is roughly proportional to the depth of U_n and inversely proportional to its cross-sectional area, and this leads to the conjecture that the limit

$$(13.2) \qquad \lim_{n\to\infty} \{n^{d-2} R_n\}$$

exists almost surely, for a large class of distributions. Similar results involving L^2 convergence have been obtained by Varadhan (1979) and Golden and Papanicolaou (1982) for the related problem of conduction in the continuous sheet $\mathbb{R}^2$, and by Künnemann (1983a, b) for the two-dimensional lattice $\mathbb{L}^2$ and edge-resistances which are uniformly bounded away from 0 and ∞. Such questions as this lie within the mathematical domain termed 'homogenization'. A great deal has been proved in the homogenization vernacular, but it remains a challenge to develop probabilistic methods for these problems. The relevant homogenization arguments have been developed by Jikov and others, and may be found in Jikov, Kozlov, and Oleinik (1994).

Subadditivity has provided a useful probabilistic technique. Whereas electrical resistance is not itself a subadditive function of the underlying volume of material, Hammersley (1988) has shown that the dissipated energy is subadditive when the problem is suitably reformulated; using this observation, he has proved the existence of the limit corresponding to (13.2) for thin slabs of (non-superconducting) material. See Dal Maso and Modica (1986) also.

The simplest special case in which percolation comes to the forefront is when each edge-resistance is either an insulator or a standard conductor, which is to say that there exists a number p such that

$$P\big(r(e) = 1\big) = p, \quad P\big(r(e) = \infty\big) = 1 - p,$$

for all edges e. If p is beneath the critical probability $p_c(d)$ of bond percolation on $\mathbb{L}^d$, there exist insufficiently many conducting edges to guarantee non-zero current, which is to say that

$$P\big(R_n = \infty \text{ for all large } n\big) = 1 \qquad \text{if } p < p_c(d).$$

The case $p > p_c(d)$ is of much greater interest. Jikov, Kozlov, and Oleinik (1994) have used arguments from homogenization to prove that

$$(13.3) \qquad n^{d-2} R_n \to \rho \qquad \text{almost surely as } n \to \infty,$$

for some constant $\rho = \rho(p, d)$ satisfying $0 < \rho < \infty$.

Various upper and lower bounds are known for $n^{d-2} R_n$, and some of the best are those of Chayes and Chayes (1986d). In seeking a lower bound for R_n, we observe that there is a dominating influence from those edges of U_n which are conducting and whose endvertices are contained in disjoint infinite conducting paths of $\mathbb{L}^d$. This leads by way of the BK inequality to the bound

$$\liminf_{n\to\infty}\{n^{d-2}R_n\} \geq \frac{1}{dp\theta(p)^2} \qquad \text{almost surely,}$$

where $\theta(p)$ is the percolation probability for bond percolation on $\mathbb{L}^d$. For a decent upper bound on R_n, we need to find a large number of edge-disjoint paths of conducting edges joining S_n to T_n, and it is here that Theorem (7.68) is useful. When $p > p_c$, the number of edge-disjoint conducting paths joining S_n to T_n has order n^{d-1}, and this implies that

$$(13.4) \qquad \limsup_{n\to\infty}\{n^{d-2}R_n\} < \infty \qquad \text{almost surely, if } p > p_c(d).$$

Random media are discussed in some detail in papers contained in Hughes and Ninham (1983). Other references to the above work include Kesten (1982, 1987e) and Jikov, Kozlov, and Oleinik (1994). Stinchcombe (1974), Grimmett and Kesten (1984b), and Grimmett (1985c) discuss random electrical networks on trees and complete graphs.

13.3 Stochastic Pin-Ball

A ball is propelled through an environment of reflecting obstacles. What can be said about its trajectory? Such a process was studied by Lorentz (1905) in a theoretical study of the motion of an electron through a field of massive particles. This question was pursued further by Ehrenfest (1959), partly under the title 'wind–tree model'. It has spawned a number of beautiful problems in probability most of which remain largely unanswered.

Randomness may be injected into the problem in either (or both) of two ways, namely in the environment of obstacles or in the rules governing the motion of the ball. Physicists are interested more by a deterministic motion within a random medium, but few rigorous facts are known about such systems.

We illustrate this topic with a concrete and well known special case. Let $0 \leq p \leq 1$. We designate each vertex x of the square lattice $\mathbb{L}^2$ a *mirror* with probability p, and a *crossing* otherwise; different vertices receive independent designations. If a vertex x is designated a mirror, we call x a *north-west (NW) mirror* with probability $\frac{1}{2}$ and a *north-east (NE) mirror* otherwise. We now position two-sided plane mirrors at the vertices of $\mathbb{L}^2$ in the manner of the designations, as illustrated in Figure 13.2. A ray of light is shone from the origin in a given direction. On reaching any crossing, it passes through undeflected. On striking a mirror, it is deflected through a right angle in the appropriate direction. The following

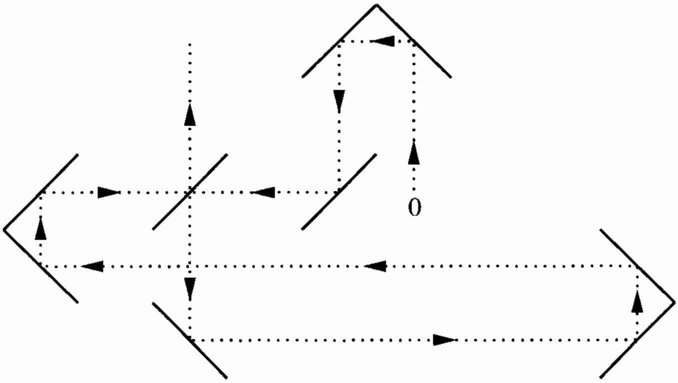


Figure 13.2. A disposition of mirrors on the square lattice. The ray of light traverses the crossings (which are not marked in this picture) and is reflected by the mirrors. The problem is to determine whether or not the light is a.s. restricted to a bounded region.

dichotomy is fairly clear. Either the light traverses an infinite path beginning at the origin (possibly containing some self-intersections), or it traverses a closed (finite) loop. We write $\xi(p)$ for the probability that the light illuminates infinitely many vertices. The principal question is to determine whether or not $\xi(p) > 0$ for a given value of p.

In this lattice version of the Lorentz model, the mirrors represent the massive particles and the light represents the electron. Percolation techniques are relevant to such lattice problems. The problem may also be formulated in the continuum $\mathbb{R}^2$; see Bunimovitch and Sinai (1981) and Spohn (1991).

It seems hard to determine whether or not $\xi(p) > 0$ for a given value of p. The heart of the difficulty lies in the fact that the mirror model constitutes a type of dynamical system in a random environment. That is to say, conditional on the environment of mirrors, the light behaves deterministically; however, its trajectory is very sensitive to small changes in the configuration of mirrors. It is trivial that $\xi(0) = 1$, and the only other fact known concerning ξ is that $\xi(1) = 0$. Numerical simulations of Ziff, Kong, and Cohen (1991) suggest that $\xi(p) = 0$ for all $p > 0$.

That $\xi(1) = 0$ is a consequence of the fact that bond percolation on $\mathbb{L}^2$ satisfies $\theta(\frac{1}{2}) = 0$. This statement we explain as follows. Let $\mathcal{L}$ be the 'diagonal lattice' having vertex set $(m + \frac{1}{2}, n + \frac{1}{2})$ for $m, n \in \mathbb{Z}$ with $m + n$ even; $\mathcal{L}$ has an edge joining $(m + \frac{1}{2}, n + \frac{1}{2})$ and $(r + \frac{1}{2}, s + \frac{1}{2})$ if and only if $|m - r| = |n - s| = 1$. The lattice $\mathcal{L}$ is drawn in Figure 13.3. We now use the mirrors of $\mathbb{L}^2$ in order to generate a bond percolation process on $\mathcal{L}$ in the following manner. An edge of $\mathcal{L}$ joining $(m - \frac{1}{2}, n - \frac{1}{2})$ to $(m + \frac{1}{2}, n + \frac{1}{2})$ is declared *open* if the vertex (m, n) of $\mathbb{L}^2$ is a NE mirror; similarly, an edge of $\mathcal{L}$ joining $(m - \frac{1}{2}, n + \frac{1}{2})$ to $(m + \frac{1}{2}, n - \frac{1}{2})$ is declared *open* if (m, n) is a NW mirror. Since $\mathcal{L}$ is isomorphic to the square lattice, the resulting percolation process is bond percolation on a square lattice at

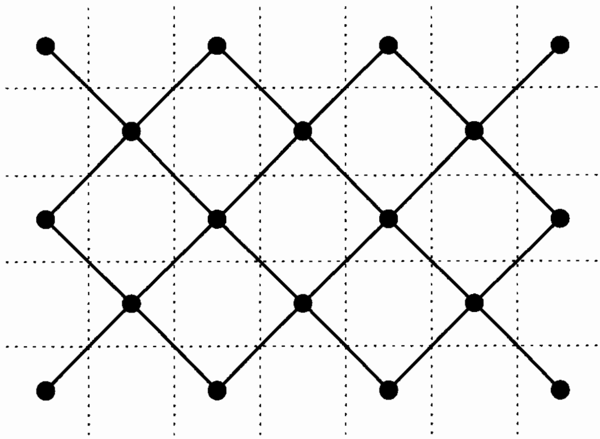


Figure 13.3. The 'diagonal' lattice $\mathcal{L}$ is obtained from $\mathbb{L}^2$ by placing vertices in the centres of every other face, and by joining such vertices by edges in the manner of this picture.

density $\frac{1}{2}p$. When $p = 1$, the density of open edges of $\mathcal{L}$ is $\frac{1}{2}$. Now, $\theta(\frac{1}{2}) = 0$ by Lemma (11.12), implying by duality that the origin of $\mathbb{L}^2$ is contained a.s. in the interior of some open circuit F of $\mathcal{L}$. Each edge of F corresponds to a parallel mirror on the corresponding vertex of the original lattice, and therefore the circuit F gives rise to a enclosure of mirrors surrounding the origin. Light cannot escape such an enclosure, and therefore $\xi(1) = 0$.

There are numerous related systems which pose interesting challenges to the physicist and mathematician. For example, there exist many other types of reflecting body in two or more dimensions than just simple plane mirrors. Very little seems to be known about the trajectories of light rays through general 'random labyrinths' of mirrors. See Grimmett (1997) for a general formulation of such a problem, as well as for further references to the argument presented above.

Ruijgrok and Cohen (1988) have proposed a study of 'rotator' models as well as 'mirror' models. In the simplest rotator model, every vertex of $\mathbb{L}^2$ is designated either a rotator or a crossing; rotators come in two types, namely 'right' rotators and 'left' rotators. As before, light traverses a crossing without deflection, but when incident on a right (respectively left) rotator, it is deflected through a right angle clockwise (respectively anticlockwise). Not a great deal has been proved about rotator models. Some results and further discussion may be found in Bunimovitch and Troubetzkoy (1992, 1994).

In another variant of such systems, a little extra randomness is introduced into the environment in the following manner. Let p_{rw} be a real number satisfying $0 < p_{\text{rw}} < 1$. We designate each vertex a 'random-walk point' with probability p_{rw}; vertices which are not random-walk points may be mirrors or crossings as above. When light strikes a random-walk point, its exit direction is chosen uniformly at random, independently of the environment and of all earlier choices; in dimension

d, each of the $2d$ possible directions has equal probability $(2d)^{-1}$. Percolation methods are useful for describing the geometry of the ensuing labyrinth. Here are brief statements of two partial results.

Consider a general reflecting labyrinth on the d-dimensional cubic lattice $\mathbb{L}^d$, having a strictly positive density p_{rw} of random-walk points; we assume that $d \geq 2$. The states of different vertices are assumed independent of one another, and each vertex may be a crossing, a random-walk point, or one of a given family of reflecting bodies. If either $p_{\text{rw}} > p_c^{\text{site}}(\mathbb{L}^d)$, or the density of non-trivial reflectors is sufficiently small (a reflector is called non-trivial if it is not a crossing),

(i) the light illuminates an infinite set with strictly positive probability; and

(ii) on the latter event, the position of the light after n steps has (asymptotically) the normal distribution with mean 0 and variance δn.

The diffusion constant δ is strictly positive, and its numerical value depends on the parameters of the labyrinth of mirrors.

Such results may be obtained by applying percolation techniques to a certain 'block' process, similar to the static renormalization of Section 7.4. This general approach is of wide applicability, and illustrates well some of the contemporary value of percolation theory. One begins by partitioning the lattice into large blocks. Associated with each block B is a certain event E_B, which is constructed with two targets in mind:

(a) the probability of each E_B may be made close to 1 by suitable choice of parameter values and block size;

(b) a cluster of blocks B for which the corresponding events E_B occur corresponds in the original lattice to certain long light paths.

Using (a), one may obtain a block process which dominates a supercritical site percolation model. Using percolation techniques and (b), one may then understand something of the geometry of the original problem. In particular, one may prove (i)–(ii) above by utilizing a suitable definition of the events E_B.

Further details may be found in Bezuidenhout and Grimmett (1997), Grimmett (1997, 1999b), and Grimmett, Menshikov, and Volkov (1996).

13.4 Fractal Percolation

Many so called 'fractals' are generated by iterative schemes, of which the classical middle-third Cantor construction is a canonical example. When the scheme incorporates a randomized step, then the ensuing set may be termed a 'random fractal'. Such sets have been studied in some generality. The following simple example is directed at a 'percolative' property, namely the possible existence in a random fractal of long paths.

We begin with the unit square $C_0 = [0, 1]^2$. At the first stage, we divide C_0 into nine (topologically closed) subsquares of side-length $\frac{1}{3}$ (in the natural way), and

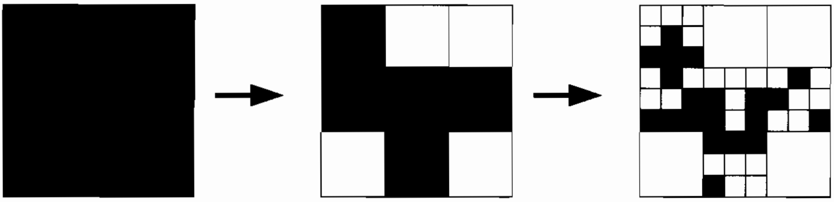


Figure 13.4. Three stages in the construction of the 'random Cantor set' C. At every stage, each retained square is replaced by a 3×3 grid of smaller squares, each of which is retained with probability p.

we declare each of the subsquares to be *open* with probability p (independently of all other subsquares). Write C_1 for the union of the open subsquares thus obtained. We now iterate this construction on each subsquare in C_1, obtaining a collection of open (sub)subsquares of side-length $\frac{1}{9}$. After k steps, we have obtained a union C_k of open squares of side-length $(\frac{1}{3})^k$. The limit set

(13.5)
$$C = \lim_{k\to\infty} C_k = \bigcap_{k\geq 1} C_k$$

is a random set whose metrical properties we wish to study. See Figure 13.4.

Constructions of the above type were introduced by Mandelbrot (1983) and studied initially by Chayes, Chayes, and Durrett (1988). Recent papers include Dekking and Meester (1990), Falconer and Grimmett (1992), Chayes (1995a, b, 1996), and Orzechowski (1996, 1998). Many generalizations of the above present themselves, for example:

(a) instead of working to base 3, we may work to base M where $M \geq 2$,

(b) replace *two* dimensions by d dimensions where $d \geq 2$,

(c) generalize the use of a square.

In what follows, (a) and (b) are generally feasible, while (c) poses a different circle of problems.

It is easily seen that the number X_k of squares present in C_k constitutes a branching process having family-size generating function $G(x) = (1 - p + px)^9$. Its extinction probability η is a root of the equation $\eta = G(\eta)$, and is such that

$$\eta = P_p(\text{extinction}) \begin{cases} = 1 & \text{if } p \leq \frac{1}{9}, \\ < 1 & \text{if } p > \frac{1}{9}. \end{cases}$$

Therefore,

(13.6)
$$P_p(C = \varnothing) = 1 \quad \text{if and only if} \quad p \leq \tfrac{1}{9}.$$

When $p > \frac{1}{9}$, the limit set C (when non-empty) is large but ramified.

(13.7) Theorem. *Let $p > \frac{1}{9}$. The Hausdorff dimension of C, conditioned on the event $\{C \neq \varnothing\}$, equals a.s.* $\log(9p)/\log 3$.

See Peyrière (1978), Hawkes (1981), Falconer (1986), Mauldin and Williams (1986), Graf (1987). Indeed, the exact Hausdorff measure function of C may be ascertained, and is found to be $h(t) = t^d(\log|\log t|)^{1-\frac{1}{2}d}$ where d is the Hausdorff dimension of C. See Graf, Mauldin, and Williams (1987).

Can C contain long paths? More concretely, can C contain a crossing from left to right of the original unit square C_0 (which is to say that C contains a connected subset having non-trivial intersection with the left and right sides of the unit square)? Let LR denote the event that such a crossing exists in C. We define the percolation probability

$$\theta(p) = P_p(\text{LR}), \tag{13.8}$$

and the critical probability

$$p_c = \sup\{p : \theta(p) = 0\}.$$

The following was proved by Chayes, Chayes, and Durrett (1988); see also Dekking and Meester (1990).

(13.9) Theorem. *We have that* $0 < p_c < 1$, *and furthermore* $\theta(p_c) > 0$.

Consider now the more general setting in which the random step involves replacing a typical square of side-length M^{-k} by a $M \times M$ grid of subsquares of common side-length $M^{-(k+1)}$ (the above concerns the case $M = 3$). For general M, a version of the above argument yields that the corresponding critical probability $p_c(M)$ satisfies $p_c(M) \geq M^{-2}$ and also

$$p_c(M) < 1 \quad \text{if } M \geq 3. \tag{13.10}$$

When $M = 2$, we need a special argument in order to obtain that $p_c(2) < 1$, and this may be achieved by using the following coupling of the cases $M = 2$ and $M = 4$. Divide C_0 into a 4×4 grid and do as follows. At the first stage, with probability p we retain all four squares in the top left corner, otherwise we delete all four; we do similarly for the three batches of four squares in each of the other three corners of C_0. Now for the second stage: examine each subsquare of side-length $\frac{1}{4}$ so far retained, and delete such a subsquare with probability $1 - p$ (different subsquares being treated independently). Note that the probability measure at the first stage dominates (stochastically) product measure with intensity π so long as $(1-\pi)^4 \geq 1 - p$. Choose π to satisfy equality here. The composite construction outlined above dominates (stochastically) a single step of a 4×4 random fractal with parameter $p\pi = p\big(1 - (1-p)^{\frac{1}{4}}\big)$, which implies that

$$p_c(2)\big(1 - (1 - p_c(2))^{\frac{1}{4}}\big) \leq p_c(4)$$

and therefore $p_c(2) < 1$ by (13.10).

Random fractals have many phases, of which the existence of left–right crossings characterizes only one. A weaker property than the existence of crossings is that the projection of C onto the x-axis is the whole interval $[0, 1]$. Projections of random fractals are of independent interest (see, for example, the 'digital sundial' theorem of Falconer (1986)). Dekking and Meester (1990) have cast such properties within a more general morphology.

We write C for a random fractal in $[0, 1]^2$. The projection of C is denoted as

$$\pi C = \{x \in \mathbb{R} : (x, y) \in C \text{ for some } y\},$$

and λ denotes Lebesgue measure. We say that C lies in one of the following phases if it has the stated property. A set is said to *percolate* if it contains a left–right crossing of $[0, 1]^2$; *dimension* is denoted by 'dim'.

I. $C = \varnothing$ a.s.

II. $P(C \neq \varnothing) > 0$, $\dim(\pi C) = \dim(C)$ a.s.

III. $\dim(\pi C) < \dim(C)$ a.s. on the event $\{C \neq \varnothing\}$, and $\lambda(\pi C) = 0$ a.s.

IV. $0 < \lambda(\pi C) < 1$ a.s. on the event $\{C \neq \varnothing\}$.

V. $P(\lambda(\pi C) = 1) > 0$ but C does not percolate a.s.

VI. $P(C \text{ percolates}) > 0$.

In many cases of interest, there is a parameter p, and the ensuing fractal moves through the phases, from I to VI, as p increases from 0 to 1. There may be critical values $p_{M,N}$ at which the model moves from Phase M to Phase N. In a variety of cases, the critical values $p_{\mathrm{I,II}}$, $p_{\mathrm{II,III}}$, $p_{\mathrm{III,IV}}$ can be determined exactly, whereas $p_{\mathrm{IV,V}}$ and $p_{\mathrm{V,VI}}$ can be much harder to find.

Here is a reasonably large class of random fractals. As before, they are constructed by dividing a square into 9 equal subsquares. In this more general system, we are provided with a probability measure μ, and we replace a square by the union of a random collection of subsquares sampled according to μ. This process is iterated on all relevant scales. Certain parameters are especially relevant. Let σ_l be the number of subsquares retained from the lth column, and let $m_l = E(\sigma_l)$ be its mean.

(13.11) Theorem. *We have that*:

(a) *if no σ_l equals* 1 *a.s., then $C = \varnothing$ if and only if $\sum_{l=1}^{3} m_l \le 1$*;

(b) $\dim(\pi C) = \dim(C)$ *a.s. if and only if*

$$\sum_{l=1}^{3} m_l \log m_l \le 0;$$

(c) $\lambda(\pi C) = 0$ *a.s. if and only if*

$$\sum_{l=1}^{3} \log m_l \le 0.$$

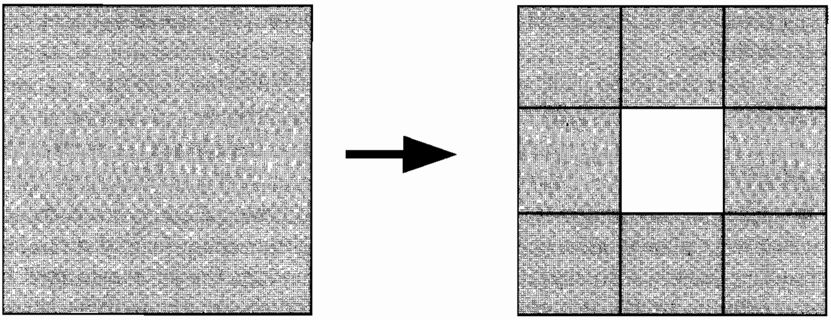


Figure 13.5. In the construction of the Sierpinski carpet, the middle square is always deleted.

It follows from Theorem (13.11) that one may check the Phases I, II, III by a knowledge of the m_l only. For the proofs, see Dekking and Grimmett (1988) and Falconer and Grimmett (1992).

Next, we apply Theorem (13.11) to the fundamental random fractal described at the beginning of this section to obtain that, for this model, $p_{\mathrm{I,II}} = \frac{1}{9}$, and we depart Phase II as p increases through the value $\frac{1}{3}$. The system is never in Phase III (by Theorem (13.11)(c)) or in Phase IV (by Theorem 1 of Falconer and Grimmett (1992)). It turns out that $p_{\mathrm{II,V}} = \frac{1}{3}$ and $\frac{1}{2} < p_{\mathrm{V,VI}} < 1$.

For the 'random Sierpinski carpet' (RSC) the picture is rather different. This set is constructed in a way similar to the above but with one crucial difference: at each iteration, the central square is removed with probability one, and the others with probability $1 - p$ (see Figure 13.5). Applying Theorem (13.11) we find that

$$p_{\mathrm{I,II}} = \tfrac{1}{8}, \quad p_{\mathrm{II,III}} = 54^{-\frac{1}{4}}, \quad p_{\mathrm{III,IV}} = 18^{-\frac{1}{3}},$$

and it happens that

$$\tfrac{1}{2} < p_{\mathrm{IV,V}} \le 0.8085, \quad 0.812 \le p_{\mathrm{V,VI}} \le 0.991.$$

Note that this process belongs to any given phase for a non-empty interval of values of p. See Dekking and Meester (1990) for more details.

Here is a conjecture which has received some attention. Writing p_c (resp. $p_c(\mathrm{RSC})$) for the critical point of the random fractal at the beginning of this section (respectively the random Sierpinski carpet), it is evident that $p_c \le p_c(\mathrm{RSC})$. Prove or disprove the strict inequality $p_c < p_c(\mathrm{RSC})$.

Peres (1996) has discovered a link between fractal percolation and Brownian Motion, via a notion called 'intersection-equivalence'. For a region $U \subseteq \mathbb{R}^d$, we call two random sets B and C *intersection-equivalent in* U if

$$(13.12) \qquad P(B \cap \Lambda \neq \varnothing) \asymp P(C \cap \Lambda \neq \varnothing) \text{ for all closed } \Lambda \subseteq U,$$

which is to say that there exist positive finite constants c_1, c_2, possibly depending on U, such that

$$c_1 \leq \frac{P(B \cap \Lambda \neq \varnothing)}{P(C \cap \Lambda \neq \varnothing)} \leq c_2$$

for all closed $\Lambda \subseteq U$.

We apply this definition for two particular random sets. First, write B for the range of Brownian Motion in $\mathbb{R}^d$, starting at a point chosen uniformly at random in the unit cube. Also, for $d \geq 3$, let C be a random fractal constructed by binary splitting (rather than the ternary splitting used above) and with parameter $p = 2^{2-d}$.

(13.13) Theorem. *Suppose that $d \geq 3$. The random sets B and C given above are intersection-equivalent.*

A similar result is valid when $d = 2$, but with a suitable redefinition of the random set C. This is achieved by taking different values of p at the different stages of the construction, namely $p = k/(k+1)$ at the kth stage.

This correspondence is not only beautiful and surprising, but also useful. It provides a fairly straightforward route to certain results concerning intersections of random walks and Brownian Motions, for example. Conversely, using the rotation invariance of Brownian Motion, one may obtain results concerning projections of the random fractal onto other directions than onto an axis (thereby complementing results of Dekking and Grimmett (1988), in the case of the special value of p given above).

The proof of Theorem (13.13) is analytical, and proceeds by utilizing:

- classical potential theory for Brownian Motion;
- the relationship between capacity and percolation for trees (Lyons (1992)); and
- the relationship between capacity on trees and capacity on an associated Euclidean space (Benjamini and Peres (1992), Pemantle and Peres (1996b)).

It is an attractive target to understand Theorem (13.13) via a coupling of the two random sets.

13.5 Contact Model

Realistic models for the spread of disease must incorporate information about the interactions between individuals, and such interactions are often governed by a spatial distribution. In 1974, Harris introduced the *contact process* as a model for a spatial epidemic, and he proved some striking results. His model may be viewed as a percolation process in which one axis, representing time, is continuous and oriented.

We suppose that individuals are positioned at the vertices of $\mathbb{L}^d$ where $d \geq 1$, and we let λ and δ be strictly positive constants. At each time t, the individual at vertex x may be in either of two possible states labelled 1 and 2; the state 1 means 'ill' or 'infected', and the state 2 means 'susceptible' (to illness). We postulate that the disease is transmitted between individuals according to the following rules. If the individual at x has state 1 (that is, is ill) at time t, this individual becomes susceptible during the short time interval $(t, t+h)$ with probability $\delta h + o(h)$. Here, δ is the *rate of cure* of the disease. If the individual at x has state 2 (that is, is susceptible) at time t, it becomes ill during the interval $(t, t+h)$ with probability $\lambda nh + o(h)$ where n is the number of ill neighbours of x. Here, λ is the *rate of infection*. Thus, cures occur spontaneously at rate δ, and infection spreads at rate λ by way of contact between infected individuals and susceptible neighbours.

The principal question is whether or not the disease survives forever. Suppose that, at time 0, all individuals are susceptible except the origin which is ill. Let

$$\psi(\lambda, \delta) = P_{\lambda,\delta}\big(\text{infection exists at all times } t \geq 0\big)$$

where $P_{\lambda,\delta}$ is the appropriate probability measure. It is not hard to see, by resetting the speed of the clock, that $\psi(\lambda, \delta)$ is a function of the ratio λ/δ only, and we therefore write $\psi(\lambda) = \psi(\lambda, 1)$. One may see by a coupling argument that $\psi(\lambda)$ is non-decreasing in λ, whence there exists a critical value λ_c such that

$$\psi(\lambda)\begin{cases} = 0 & \text{if } \lambda < \lambda_c, \\ > 0 & \text{if } \lambda > \lambda_c. \end{cases}$$

Harris (1974) proved amongst other things that λ_c is non-trivial, in the sense that $0 < \lambda_c < \infty$ if $d \geq 1$. (Note that the true dimensionality of the process in space–time is $d+1$.)

The analogy with percolation is strong, with ψ taking the role of the percolation probability θ. A good way of seeing this to the full is via the following 'graphical representation' introduced by Harris (1978). We embed the lattice $\mathbb{L}^d$ in the larger space $S = \mathbb{Z}^d \times [0, \infty)$; a typical point $s = (x, t) \in S$ represents the vertex x at time t. We now add a collection of random marks to S, whose interpretations will be clear soon. On each 'time line' $x \times [0, \infty)$, for $x \in \mathbb{Z}^d$, we place a collection of marks (which might be called 'points of cure') in the manner of a Poisson process having intensity δ. Next, for each ordered pair x, y of neighbours of $\mathbb{L}^d$, we place arrows (called 'arrows of infection') directed from x to y along the line $x \times [0, \infty)$ in the manner of a Poisson process with intensity λ. See Figure 13.6.

We interpret these marks as follows. Suppose that the vertex x is ill at time t. It remains ill until the first subsequent point of cure, which occurs at time

$$T = \inf\big\{s > t : (x, s) \text{ is a point of cure}\big\}.$$

Meanwhile, whenever there exists an arrow oriented from x to y during the time-interval $x \times (t, T]$, the infection at x spreads to y. If y is already ill, the new

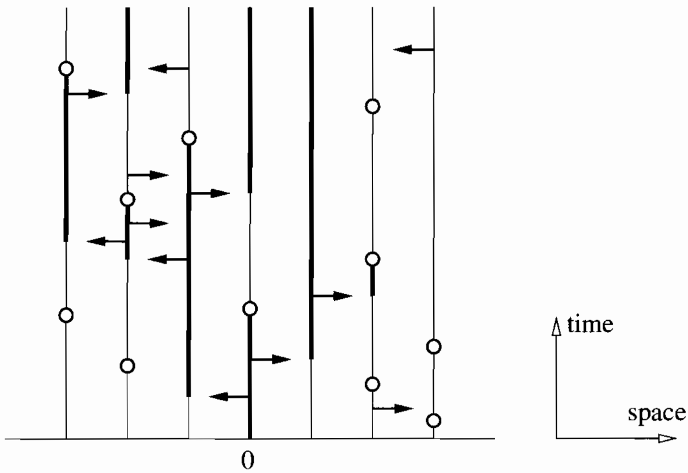


Figure 13.6. An example of the 'graphical representation' of the contact process when $d = 1$. The horizontal line represents 'space', and the vertical line above a point x represents the time axis at x. The marks $\circ$ are the points of cure, and the arrows are the arrows of infection. If we are told that, at time 0, the origin is the unique infected point, all subsequent infections may be charted by following the evolution of the graph in the direction of increasing time, and by conforming to the points of cure and the arrows of infection. In this picture, the initial infective is marked 0, and the bold lines indicate the portions of space–time which correspond to the infection of the vertex in question.

infection has no effect. It is a simple matter to check that the infection spreads in the manner of the contact process, as described above.

With infection originating at the origin only, the infection continues forever if and only if S contains an infinite path which begins at the origin, which moves in the direction of increasing time only, and which is permitted to traverse arrows of infection. The probability of this event is the percolation probability for the partly-oriented partly-continuous percolation system constructed above.

It is not surprising that percolation technology may be adapted in order to study the contact process, and vice versa. For example, one may prove by block arguments closely related to those of Chapter 7 that:

(a) the critical contact process dies out, which is to say that $\psi(\lambda_c) = 0$;

(b) if $\lambda > \lambda_c$ and infection originating at the origin continues forever, then the disease spreads in the manner of a well defined shape.

The 'shape theorem' referred to in (b) is similar to that described in Section 12.9 for first-passage percolation. See Bezuidenhout and Grimmett (1990) and Durrett (1991) for more details and references.

In more realistic epidemic models, individuals experience periods of 'removal' after being cured. Removal periods represent periods of invulnerability to infection, and can be of infinite length in the case of 'death'. For a fatal disease which invariably kills infected individuals, all removal periods are infinite, and such an 'epidemic without recovery' corresponds to the contact process with $\delta = 0$. This system is quite different from that considered above; infection never recurs, but must either become extinct, or be driven ever onwards in the manner perhaps of a fairy ring, or perhaps of the boundary of a forest fire which has consumed its interior. Kuulasmaa (1982) showed that this system also is percolative, but in the following sense. Let $\Delta_v x$ denote the set of all neighbours of a vertex x. For each x, we draw oriented edges from x to some random subset of $\Delta_v x$ chosen according to a certain probability function μ. Let $\theta(\mu)$ denote the probability that this 'partly dependent' oriented percolation model contains an infinite oriented path beginning at the origin. If μ is chosen correctly, then $\theta(\mu)$ equals the probability $\psi(\lambda, 0)$ that infection originating at the origin reaches infinitely many vertices (in the above epidemic without recovery). Even though this percolation process is not constructed entirely from independent events, some of the techniques of percolation theory may be extended in order to study its geometry, thereby gaining information about the epidemic without recovery.

There is an intermediate type of epidemic in which recovery takes place after finite time intervals. Such processes can be harder to study, since they generally lack even an elementary property of stochastic monotonicity. Recall from the graphical representation of Figure 13.6 that, for the usual contact process with $\lambda > 0$ and $\delta \geq 0$, the greater is the initial set of infectives, the more extensive is the spread of the disease. This can be false for epidemics with removal, for the following reason. By adding an extra infective, one may subsequently infect a point which, during its removal period, prevents the infection from spreading further. A forest fire may be impeded by burning a pre-emptive firebreak.

See Liggett (1985) and Durrett (1991) for further information about contact processes, and van den Berg, Grimmett, and Schinazi (1998) for recent results concerning both the contact process without recovery and a more general contact process incorporating temporary removals.

13.6 Random-Cluster Model

Probably the most famous spatial model of statistical physics is the Ising model for ferromagnetism. Founded in work of Ising (1925) and Lenz, this process has provided the setting for the development of a repertoire of techniques of wide applicability. The underlying physical phenomenon is the following. A piece of iron is placed in a magnetic field, which is increased from zero to some maximum, and then reduced back to zero. The iron may retain some residual magnetization, but only when the temperature is not too high. There exists a critical temperature T_c marking the threshold between the phase where residual magnetization is retained and the phase where none is retained. In the Ising model for the ferromagnet, the iron is modelled by a lattice each vertex of which may be in either of two states labelled $+$ and $-$, representing the 'spin' of a notional particle at that vertex. The configuration of spins is governed by a certain probability measure which favours adjacent pairs of like spins.

Potts (1952) proposed a generalization of the Ising model in which each vertex may be in any of q distinct states, labelled $1, 2, \ldots, q$; the Ising model is recovered when $q = 2$. In common with the Ising model, the Potts model also has a phase transition at some critical temperature $T_c(q)$. [Note: in this section, q is no longer defined to equal $1 - p$.]

In the late 1960s, Fortuin and Kasteleyn discovered that the Ising and Potts models, together with the percolation model, may be placed within a unified system having a coherent methodology. This discovery led to their formulation of a process which they called a 'random-cluster model' (see Fortuin (1972b) and the references therein). The construction is unusual, and arose from the observation that all these systems satisfy versions of the 'series and parallel laws'. A historical account of the origins of the random-cluster model may be found in Grimmett (1994b).

Here is the definition of the random-cluster model on a finite graph. Let $0 \le p \le 1$ and $q > 0$. Let $G = (V, E)$ be a finite graph, and let F be a subset of E chosen according to the following probability mass function $\varphi_{p,q}$:

$$\text{(13.14)} \qquad \varphi_{p,q}(F) = \frac{1}{Z} p^{|F|}(1-p)^{|E \setminus F|} q^{k(F)} \quad \text{for } F \subseteq E,$$

where $k(F)$ is the number of connected components of the graph (V, F), and Z is a normalizing constant chosen so that

$$\sum_{F \subseteq E} \varphi_{p,q}(F) = 1.$$

If $q = 1$, then

$$\varphi_{p,1}(F) = p^{|F|}(1-p)^{|E \setminus F|},$$

which is to say that different edges are present independently of one another, each with probability p; this is bond percolation on G. It turns out that, for $q = 2, 3, \ldots$, the probability distribution $\varphi_{p,q}$ is closely related to that of the Potts model with q states, and in which p is a certain function of the temperature T and 'pair-interaction' J. In the particular case $q = 2$, $\varphi_{p,2}$ corresponds to the probability measure governing the Ising model. We do not give details of this correspondence, but state the principal observation briefly. Consider a Potts model on G having q states and governed by the probability measure

$$\pi_{J,T}(\sigma) = \frac{1}{Y} \prod_{e \in E} \exp\{J\delta_\sigma(e)/T\}, \quad \text{for } \sigma \in \Sigma = \{1, 2, \ldots, q\}^V,$$

where Y is the normalizing constant, and the Kronecker delta $\delta_\sigma(e)$ is given for the edge $e = \langle u, v \rangle$ by

$$\delta_\sigma(e) = \begin{cases} 1 & \text{if } \sigma(u) = \sigma(v), \\ 0 & \text{otherwise.} \end{cases}$$

Let

$$p = 1 - e^{-J/T}.$$

It may be shown that the two-point function of the Potts model differs only by a multiplicative factor from the two-point connectivity function of the random-cluster model with parameters p and q; that is to say,

$$\pi_{J,T}(\sigma_x = \sigma_y) - \frac{1}{q} = (1 - q^{-1})\varphi_{p,q}(x \leftrightarrow y).$$

Using this correspondence, questions concerning the correlation structure of the Potts model may be transformed into questions involving the stochastic geometry of the random-cluster model. The simplest proof of this correspondence was discovered by Edwards and Sokal (1988).

Unlike the percolation model, it is not completely straightforward to define the random-cluster model on an *infinite* lattice such as $\mathbb{L}^d$; formula (13.14) simply does not work in this case. The usual approach to this problem is to let $\varphi_{\Lambda,p,q}$ be the random-cluster measure on some finite region Λ of the lattice, and to pass to the limit as Λ expands to fill out the whole space. This process is known as 'passing to the thermodynamic limit'. In order to achieve a proper definition of the infinite-volume limit, some discussion of 'boundary conditions' is necessary. We restrict ourselves here to 'wired boundary conditions', which is to say that, in counting the number $k(F)$ of clusters corresponding to an edge-subset F of Λ, we treat all clusters intersecting the boundary $\partial\Lambda$ as if they were one single cluster. When this is correctly formulated, it turns out that the limit measure

$$\varphi_{p,q} = \lim_{\Lambda \to \mathbb{Z}^d} \varphi_{\Lambda,p,q}$$

exists when $q \geq 1$, and satisfies the FKG inequality. Let

$$\theta(p, q) = \varphi_{p,q}\big(0 \text{ belongs to an infinite cluster}\big)$$

denote the corresponding percolation probability. For $q \geq 1$, there exists a critical value of p, denoted $p_c(q)$, such that

$$\theta(p, q)\begin{cases} = 0 & \text{if } p < p_c(q), \\ > 0 & \text{if } p > p_c(q). \end{cases}$$

Furthermore, $p_c(1)$ is the critical probability of bond percolation, and, when $q = 2, 3, \dots$, the critical value $p_c(q)$ is related to the critical temperature $T_c(q)$ of the q-state Potts model by

$$p_c(q) = 1 - e^{-J/T_c(q)}.$$

In this sense (and indeed further) the phase transition of the random-cluster model generalizes those of percolation, Ising, and Potts models.

A body of techniques for the random-cluster model has emerged in recent years, but many central questions remain unanswered. For example, two known facts when $q \geq 1$ and $d \geq 2$ are:

(i) if $\theta(p, q) > 0$, there exists a.s. a *unique* infinite cluster;

(ii) if q is large, say $q > Q = Q(d)$, then $\theta(\cdot, q)$ is discontinuous at the critical point, which is to say that $\theta(p_c(q), q) > 0$.

See Kotecký and Shlosman (1982), Laanait, Messager, Miracle-Sole, Ruiz, and Shlosman (1991), and Grimmett (1995b, 1997). Fact (ii) is particularly striking, since it contrasts with the corresponding conjecture of continuity for percolation.

One may conjecture that

(iii) if q is small, say $1 \leq q < Q(d)$, then $\theta(p_c(q), q) = 0$;

(iv) if $q \geq 1$ and $p < p_c(q)$, the probability $\varphi_{p,q}(\text{rad}(C) \geq n)$ decays exponentially as $n \to \infty$.

There is a rich family of conjectures for random-cluster models, ranging from exact calculations, conformal invariance, and a Cardy formula when $d = 2$ and $1 \leq q < 4$, to the belief that $Q(2) = 4$ and $Q(d) = 2$ when $d \geq 6$. In addition, very little is known when $0 < q < 1$.

Perhaps the most stimulating conjecture is the following prediction for the value of the critical point of the random-cluster model on the two-dimensional square lattice $\mathbb{L}^2$: an argument utilizing duality suggests that

$$p_c(q) = \frac{\sqrt{q}}{1 + \sqrt{q}} \quad \text{for } q \geq 1.$$

This exact value is known to be valid when $q = 1$ (by Theorem 11.11), when $q = 2$ (by exact calculations valid for the Ising model), and for $q \geq 25.72$ (see the references in Grimmett (1997)).

This beautiful model is an outstanding challenge to mathematicians. It promises a unified structure which will explain further the Ising and Potts models in the context of percolation. It indicates a mechanism for moving between models which will find further applications in statistical physics, and via methods of Monte Carlo simulation to statistical science. Further accounts include Aizenman, Chayes, Chayes, and Newman (1988), Baxter (1982), Grimmett (1995b, 1997), Grimmett and Piza (1997), and Häggström (1998).

Appendix I

The Infinite-Volume Limit for Percolation

Certain steps were omitted from the proof of Theorem (5.48). The calculations necessary for these minor steps are rather standard, and we indicate here how they may be performed.

We shall use the notation of Section 5.3 without further elaboration. The principal quantities are

$$\theta(p,\gamma) = 1 - \sum_{n=1}^{\infty}(1-\gamma)^n P_p(|C| = n), \tag{I.1}$$

together with the corresponding quantity defined on the box $B(N)$ with periodic boundary conditions,

$$\theta_N(p,\gamma) = 1 - \sum_{n=1}^{\infty}(1-\gamma)^n P_p(|C_N| = n), \tag{I.2}$$

where C_N is the open cluster of $B(N)$ containing the origin.

We are required to prove the following:

$$\theta(p,\gamma) \text{ is continuously differentiable in } p, \qquad \text{for } \gamma > 0, \tag{I.3}$$

and

$$\theta_N(p,\gamma) \to \theta(p,\gamma) \qquad \text{as } N \to \infty, \tag{I.4}$$

$$\frac{\partial \theta_N}{\partial p} \to \frac{\partial \theta}{\partial p} \qquad \text{as } N \to \infty, \tag{I.5}$$

$$\frac{\partial \theta_N}{\partial \gamma} \to \frac{\partial \theta}{\partial \gamma} \qquad \text{as } N \to \infty, \tag{I.6}$$

if $0 < p, \gamma < 1$. In demonstrating these limits we use techniques related to those of Chapter 4 and Section 8.7. Thus we write

$$P_p(|C| = n) = \sum_{m,b} a_{nmb} p^m (1-p)^b \tag{I.7}$$

and

$$P_p(|C_N| = n) = \sum_{m,b} a_{nmb}(N) p^m (1-p)^b, \tag{I.8}$$

where a_{nmb} (respectively $a_{nmb}(N)$) is the number of animals of $\mathbb{L}^d$ (respectively $B(N)$ with periodic boundary conditions) having n vertices, m edges, and b boundary edges. We note as in (4.14) and (4.15) that $a_{nmb} = a_{nmb}(N) = 0$ if either $b > 2dn$ or $m > dn$, and we may therefore assume that $b \le 2dn$ and $m \le dn$.

We prove (I.3) first, restricting ourselves to the case when $0 < p < 1$; the special cases $p = 0, 1$ are of no importance in Section 5.3 and may be dealt with somewhat similarly. Suppose then that $0 < p < 1$ and $\gamma > 0$. It suffices to show that the term-by-term derivative of the infinite series in (I.1) is uniformly convergent on some neighbourhood of p. The tail of the term-by-term derivative is

$$\sum_{n=N}^{\infty} (1-\gamma)^n \sum_{m,b} a_{nmb} p^m (1-p)^b \left(\frac{m}{p} - \frac{b}{1-p} \right),$$

where we have used (I.7). This is no larger in absolute value than

$$\sum_{n=N}^{\infty} (1-\gamma)^n \frac{2dn}{p(1-p)} P_p(|C| = n)$$

since $m, b \le 2dn$; the last summation converges uniformly as $N \to \infty$ for values of p in any strict sub-interval of $[0, 1]$. It follows that $\theta(\cdot, \gamma)$ is continuously differentiable at p.

We move on to (I.4)–(I.6) of which we shall prove (I.4) only, the other two limits having analogous proofs. We have from (I.1) and (I.2) that

$$\left|\theta(p,\gamma) - \theta_N(p,\gamma)\right| \le \sum_{n=1}^{\infty} (1-\gamma)^n \left| P_p(|C| = n) - P_p(|C_N| = n) \right|.$$

Now $P_p(|C| = n) = P_p(|C_N| = n)$ if $n < N$, since no cluster of $\mathbb{L}^d$ containing the origin intersects the boundary of $B(N)$ unless it has more than N vertices. Therefore,

$$\begin{aligned} \left|\theta(p,\gamma) - \theta_N(p,\gamma)\right| &\le \sum_{n=N}^{\infty} 2(1-\gamma)^n \\ &= 2(1-\gamma)^N/\gamma \qquad \text{if } \gamma > 0, \end{aligned}$$

and hence $\theta_N(p,\gamma) \to \theta(p,\gamma)$ as $N \to \infty$ whenever $\gamma > 0$. The remaining limits (I.5) and (I.6) may be proved by using the expansion (I.8) and combining the arguments leading to (I.3) and (I.4).

Appendix II

The Subadditive Inequality

Let $(x_r : r \geq 1)$ be a sequence of real numbers. If the x_r satisfy the 'subadditive inequality'

(II.1) $$x_{m+n} \leq x_m + x_n \qquad \text{for all } m, n,$$

we say that the sequence is *subadditive*. Subadditive sequences are nearly additive, in the sense of the following theorem.

(II.2) Theorem. Subadditive limit theorem. *If $(x_r : r \geq 1)$ is subadditive, the limit*

(II.3) $$\lambda = \lim_{r \to \infty} \left\{ \frac{1}{r} x_r \right\}$$

exists and satisfies $-\infty \leq \lambda < \infty$. Furthermore,

(II.4) $$\lambda = \inf \left\{ \frac{1}{m} x_m : m \geq 1 \right\},$$

and thus $x_m \geq m\lambda$ for all m.

See Hille (1948, Theorem 6.6.1) for a proof of this standard result. In many potential applications of the subadditive limit theorem, we encounter sequences $(x_r : r \geq 1)$ which are nearly subadditive, in the sense that they satisfy a 'generalized subadditive inequality' of the form

(II.5) $$x_{m+n} \leq x_m + x_n + g_{m+n} \qquad \text{for all } m, n,$$

for some given sequence $(g_r : r \geq 1)$. Such a sequence $(x_r : r \geq 1)$ is nearly additive, so long as the g_r do not grow too fast.

(II.6) Theorem. Generalized subadditive limit theorem. *Suppose that the sequence* $(x_r : r \geq 1)$ *satisfies the generalized subadditive inequality* (II.5), *where* $(g_r : r \geq 1)$ *is a non-decreasing sequence satisfying*

(II.7)
$$\sum_{r=1}^{\infty} \frac{g(r)}{r(r+1)} < \infty.$$

The limit

(II.8)
$$\lambda = \lim_{r\to\infty} \left\{ \frac{1}{r} x_r \right\}$$

exists and satisfies $-\infty \leq \lambda < \infty$. *Furthermore,*

(II.9)
$$x_m \geq m\lambda + g_m - 4m \sum_{r=2m}^{\infty} \frac{g(r)}{r(r+1)} \qquad \textit{for all } m.$$

This was proved by Hammersley (1962). A similar result was found by M. E. Fisher (unpublished) in independent and roughly contemporaneous work.

It is not unusual to encounter sequences which satisfy inequalities somewhat similar to the generalized subadditive inequality; for example, in Chapter 8 we came across a real sequence $(x_r : r \geq 1)$ satisfying

(II.10)
$$x_{m+n+2} \leq x_m + x_n + g_n \qquad \text{for all } m, n.$$

It is not difficult to adapt the proof of the subadditive limit theorem to deal with such inequalities. It turns out that (II.10), together with the condition $n^{-1} g_n \to 0$ as $n \to \infty$, is sufficient to ensure the existence of $\lambda = \lim_{r\to\infty}\{r^{-1} x_r\}$, and one obtains also that

(II.11)
$$x_r \geq (r+2)\lambda - g_r \qquad \text{for all } r.$$

It is easy to obtain more refined versions of such results.

List of Notation

Against each entry appears the number of the section in which the notation first appears.

Graphs and sets:

G_{c}	1.6	Covering graph of G
G_{d}	11.2	Dual graph of G
$\langle a, b \rangle$	1.3	Edge joining vertices a and b
$\mathbb{Z}$	1.3	The set $\{\dots, -2, -1, 0, 1, 2, \dots\}$ of integers
$\mathbb{Z}_+$	7.3	The set $\{0, 1, 2, \dots\}$ of non-negative integers
$\mathbb{L}^d$	1.3	The d-dimensional cubic lattice
$\mathbb{T}$	3.1	The triangular lattice
$\mathbb{H}$	11.9	The hexagonal lattice
	7.3	The half-space $\mathbb{Z}^{d-1} \times \mathbb{Z}_+$
$\mathbb{E}^d$	1.3	The set of edges of $\mathbb{L}^d$
$\mathbb{L}^2_{\mathrm{d}}$	11.2	The dual lattice of $\mathbb{L}^2$
$\mathbb{C}$	11.10	The complex plane
x_i	1.3	The value of the ith component of the vertex x of $\mathbb{L}^d$, unless otherwise specified
$B(n, x)$	1.3	Box with side-length $2n$ and centre at x
$B(n)$	1.3	Box with side-length $2n$ and centre at 0
$B(a, b)$	1.3	Box with vertices x satisfying $a_i \le x_i \le b_i$ for $i = 1, 2, \dots, d$
$A(k)$	11.7	The annulus $B(3k) \backslash B(k)$
S_k	3.3	The slab $\mathbb{Z}^2 \times \{0, 1, 2, \dots, k\}^{d-2}$
∂A	1.3	The surface of the set A of vertices
D_{e}	4.2	The set of edges of a graph D
ΔD	4.2	The edge boundary of D, being the set of edges not in D which are incident to at least one vertex of D
$\Delta_{\mathrm{v}} V$	7.2	Exterior vertex boundary of a set V
$\lambda(d)$	1.4	The connective constant of $\mathbb{L}^d$
$\mathcal{R}_{\mathrm{b}}(\mathbb{L}^d)$	8.6, 9.3	The lattice $\mathbb{L}^d$ renormalized by a linear factor b
$\delta(x, y)$	1.3	The number of edges in the shortest path from x to y

$\lvert x \rvert$	1.3	$\delta(0, x)$, for $x \in \mathbb{Z}^d$
$\lVert x \rVert$	1.3	$\max\{\lvert x_i \rvert : 1 \le i \le d\}$, for $x \in \mathbb{Z}^d$
$x \sim y$	1.3	x is adjacent to y
$\mathrm{rad}(D)$	6.1	The radius $\max\{\lvert x \rvert : x \in D\}$ of a subgraph D of $\mathbb{L}^d$ containing 0
$\mathrm{diam}(D)$	7.4	$\max\{\lvert x_i - y_i \rvert : x, y \in D,\ 1 \le i \le d\}$, the maximum of the differences of the coordinates of vertices in the graph D
$S(n, x)$	5.2	The sphere $\{y \in \mathbb{Z}^d : \lvert y - x \rvert \le n\}$ having radius n and centre at x
$S(n)$	5.2	The sphere $S(n, 0)$
	11.3	The rectangle $[0, n+1] \times [0, n]$ in $\mathbb{L}^2$
$\lvert A \rvert$	1.3	The cardinality of a set A, or the number of vertices of a graph A
a_{nmb}	4.2	Number of animals of $\mathbb{L}^d$ having n vertices, m edges, and b edges in the edge boundary

Probability notation:

p	1.3	The probability that an edge is open
q	1.3	$1 - p$
P_p	1.3	Product measure with density p
E_p	1.3	Expectation operator corresponding to P_p
P_p^f	1.3	Product measure on $\mathbb{L}^d$ with the edge f deleted
I_A	1.3	Indicator function of an event A
$E_p(X; A)$	1.3	Expectation $E_p(X I_A)$ of X on the event A
cov_p	2.5	Covariance corresponding to P_p
var_p	2.5	Variance corresponding to P_p
ω	1.3	Typical realization of open and closed edges
$K(\omega)$	1.3	The set of edges which are open in ω
$\overline{A}$, A^{c}	1.3	Complement of event A
$A \circ B$	2.3	Event that A and B occur 'disjointly', for increasing events A, B
$A \,\square\, B$	2.3	Event that A and B occur 'disjointly'
$\le_{\mathrm{st}}$	1.6	Stochastic domination inequality

Percolation notation:

$C(x)$	1.3	Open cluster at x
C	1.3	Open cluster at 0
$p_{\mathrm{c}},\ p_{\mathrm{c}}(d)$	1.4	Critical probability of bond percolation on $\mathbb{L}^d$
$p_{\mathrm{c}}(G)$, $p_{\mathrm{c}}^{\mathrm{bond}}(G)$	1.6	Critical probability of bond percolation on a connected graph G
$p_{\mathrm{c}}^{\mathrm{site}}(G)$	1.6	Critical probability of site percolation on a connected graph G
$p_{\mathrm{c}}^{\mathrm{slab}}$	7.1	Limit of slab critical probabilities

$\theta(p)$, $\theta^{\text{bond}}(p)$	1.4, 1.6	Bond percolation probability $P_p(\|C\| = \infty)$
$\theta^{\text{site}}(p)$	1.6	Site percolation probability
$\chi(p)$	1.5	Mean cluster size $E_p\|C\|$
$\chi^{\text{f}}(p)$	1.5	Mean size $E_p(\|C\|; \|C\| < \infty)$ of a finite open cluster
$\kappa(p)$	1.5	Number $E_p(\|C\|^{-1})$ of open clusters per vertex
$\xi(p)$	6.2, 8.5	Correlation length
$\tau_p(x, y)$	6.1	Connectivity function $P_p(x \leftrightarrow y)$
$\tau_p^{\text{f}}(x, y)$	8.1	Truncated connectivity function $P_p(x \leftrightarrow y,\ \|C(x)\| < \infty)$
$\alpha, \beta, \gamma, \delta$, Δ, η, ρ, ν	9.1	Critical exponents, unless otherwise specified
$\{A \leftrightarrow B\}$	1.3	Event that some vertex of A is joined by an open path to some vertex of B
$\{A \not\leftrightarrow B\}$	1.3	Complement of $\{A \leftrightarrow B\}$

Asymptotics:

$a_n \sim b_n$	1.3	$a_n/b_n \to 1$ as $n \to \infty$
$a_n \approx b_n$	1.3	$\log a_n / \log b_n \to 1$ as $n \to \infty$
$a_n \simeq b_n$	9.1	Asymptotic properties of a_n are comparable to those of b_n
$f(p) \asymp g(p)$	10.3	$f(p)/g(p)$ is bounded away from 0 and ∞ on a neighbourhood of p_c

Similar asymptotic relations are used for functions of p and n, in the limits as p approaches p_c and $n \to \infty$.

Finally:

$\lfloor c \rfloor$	1.3	Greatest integer not greater than c
$\lceil c \rceil$	1.3	Least integer not less than c

References

Ahlswede, R., Daykin, D. E.
1979 Inequalities for a pair of maps $S \times S \to S$ with S a finite set, *Mathematische Zeitschrift* 165, 267–289.

Aizenman, M.
1982 Geometric analysis of ϕ^4 fields and Ising models, *Communications in Mathematical Physics* 86, 1–48.
1987 General results in percolation theory, *Proceedings of Taniguchi Workshop and Symposium on Probabilistic Methods in Mathematical Physics* (N. Ikeda, ed.), Kataka and Kyoto, pp. 1–22.
1995 The geometry of critical percolation and conformal invariance, *Proceedings STATPHYS 1995 (Xianmen)* (Hao Bai-lin, ed.), World Scientific, pp. 104–120.
1997 On the number of incipient spanning clusters, *Nuclear Physics B* 485, 551–582.
1998 Scaling limit for the incipient spanning clusters, *Mathematics of Multiscale Materials* (K. Golden, G. Grimmett, R. James, G. Milton, P. Sen, ed.), IMA Volumes in *Mathematics and its Applications*, vol. 99, Springer, New York, pp. 1–24.

Aizenman, M., Barsky, D. J.
1987 Sharpness of the phase transition in percolation models, *Communications in Mathematical Physics* 108, 489–526.
1991 Percolation critical exponents under the triangle condition, *Annals of Probability* 19, 1520–1536.

Aizenman, M., Barsky, D. J., Fernández, R.
1987 The phase transition in a general class of Ising-type models is sharp, *Journal of Statistical Physics* 47, 343–374.

Aizenman, M., Chayes, J. T., Chayes, L., Fröhlich, J., Russo, L.
1983 On a sharp transition from area law to perimeter law in a system of random surfaces, *Communications in Mathematical Physics* 92, 19–69.

Aizenman, M., Chayes, J. T., Chayes, L., Newman, C. M.
1988 Discontinuity of the magnetization in one-dimensional $1/|x-y|^2$ Ising and Potts models, *Journal of Statistical Physics* 50, 1–40.

Aizenman, M., Delyon, F., Souillard, B.
1980 Lower bounds on the cluster size distributions, *Journal of Statistical Physics* 23, 267–280.

Aizenman, M., Fernández, R.
1986 On the critical behavior of the magnetization in high-dimensional Ising models, *Journal of Statistical Physics* 44, 393–454.
1988 Critical exponents for long-range interactions, *Letters in Mathematical Physics* 16, 39.

Aizenman, M., Grimmett, G. R.
1991 Strict monotonicity for critical points in percolation and ferromagnetic models, *Journal of Statistical Physics* 63, 817–835.

Aizenman, M., Kesten, H., Newman, C. M.
1987a Uniqueness of the infinite cluster and continuity of connectivity functions for short- and long-range percolation, *Communications in Mathematical Physics* 111, 505–532.
1987b Uniqueness of the infinite cluster and related results in percolation, *Percolation Theory and Ergodic Theory of Infinite Particle Systems* (H. Kesten, ed.), IMA Volumes in *Mathematics and its Applications*, vol. 8, Springer, New York, pp. 13–20.

Aizenman, M., Lebowitz, J.
1988 Metastability in bootstrap percolation, *Journal of Physics A: Mathematical and General* 21, 3801–3813.

Aizenman, M., Newman, C. M.
1984 Tree graph inequalities and critical behavior in percolation models, *Journal of Statistical Physics* 36, 107–143.
1986 Discontinuity of the percolation density in one-dimensional $1/|x-y|^2$ percolation models, *Communications in Mathematical Physics* 107, 611–647.

Alexander, K. S.
1990 Lower bounds on the connectivity function in all directions for Bernoulli percolation in two and three dimensions, *Annals of Probability* 18, 1547–1562.
1992 Stability of the Wulff construction and fluctuations in shape for large finite clusters in two-dimensional percolation, *Probability Theory and Related Fields* 91, 507–532.
1993a Finite clusters in high-density continuous percolation: compression and sphericality, *Probability Theory and Related Fields* 97, 35–63.
1993b A note on some rates of convergence in first-passage percolation, *Annals of Applied Probability* 3, 81–90.
1995a Simultaneous uniqueness of infinite graphs in stationary random labeled graphs, *Communications in Mathematical Physics* 168, 39–55.
1995b Percolation and minimal spanning forests in infinite graphs, *Annals of Probability* 23, 87–104.
1996 The RSW theorem for continuum percolation and the CLT for Euclidean minimal spanning trees, *Annals of Applied Probability* 6, 466–494.
1997 Approximation of subadditive functions and convergence rates in limiting-shape results, *Annals of Probability* 25, 30–55.

Alexander, K. S., Chayes, J. T., Chayes, L.
1990 The Wulff construction and asymptotics of the finite cluster distribution for two-dimensional Bernoulli percolation, *Communications in Mathematical Physics* 131, 1–50.

Andjel, E.
1993 Characteristic exponents for two-dimensional bootstrap percolation, *Annals of Probability* 21, 926–935.

Antal, P. and Pisztora, A.
1996 On the chemical distance for supercritical Bernoulli percolation, *Annals of Probability* 24, 1036–1048.

Appel, M. J., Wierman, J. C.
1987 On the absence of AB percolation in bipartite graphs, *Journal of Physics A: Mathematical and General* 20, 2527–2531.
1993 AB percolation on bond decorated graphs, *Journal of Applied Probability* 30, 153–166.

Asmussen, S., Hering, H.
1983 *Branching Processes*, Birkhäuser, Boston.

Athreya, K. B., Ney, P. E.
1972 *Branching Processes*, Springer, Berlin.

Barlow, R. N., Proschan, F.
1965 *Mathematical Theory of Reliability*, Wiley, New York.
1975 *Statistical Theory of Reliability and Life Testing*, Holt, Rinehart and Winston, New York.

Barsky, D. J., Aizenman, M.
1991 Percolation critical exponents under the triangle condition, *Annals of Probability* 19, 1520–1536.

Barsky, D. J., Grimmett, G. R., Newman, C. M.
1991a Dynamic renormalization and continuity of the percolation transition in orthants, *Spatial Stochastic Processes* (K. S. Alexander and J. C. Watkins, eds.), Birkhäuser, Boston, pp. 37–55.
1991b Percolation in half spaces: equality of critical probabilities and continuity of the percolation probability, *Probability Theory and Related Fields* 90, 111–148.

Barsky, D. J., Wu, C. C.
1998 Critical exponents for the contact process under the triangle condition, *Journal of Statistical Physics* 91, 95–124.

Batty, C. J. K., Bollman, H. W.
1980 Generalized Holley–Preston inequalities on measure spaces and their products, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 53, 157–173.

Baxter, R. J.
1982 *Exactly Solved Models in Statistical Mechanics*, Academic Press, London.

Beijeren, H. van
1982 Transport properties of stochastic Lorentz models, *Reviews in Modern Physics* 54, 195–234.

Ben-Avraham, D., Havlin, S.
1982 Diffusion on percolation clusters at criticality, *Journal of Physics A: Mathematical and General* 15, L691–L697.

Benjamini, I.
1996 Percolation on groups, unpublished manuscript.

Benjamini, I., Kalai, G., Schramm, O.
1998 Noise sensitivity of Boolean functions and applications to percolation (to appear).

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

Benjamini, I., Lyons, R., Peres, Y., Schramm, O.
1997 Group-invariant percolation on graphs, *Geometric and Functional Analysis* (to appear).
1998 Critical percolation on any nonamenable group has no infinite clusters, *Annals of Probability* (to appear).

Benjamini, I., Pemantle, R., Peres, Y.
1998 Unpredictable paths and percolation, *Annals of Probability* 26, 1198–1211.

Benjamini, I., Peres, Y.
1992 Random walks on a tree and capacity in the interval, *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* 28, 557–592.
1994 Tree-indexed random walks on groups and first passage percolation, *Probability Theory and Related Fields* 98, 91–112.

Benjamini, I., Schramm, O.
1996 Percolation beyond $\mathbb{Z}^d$, many questions and a few answers, *Electronic Communications in Probability* 1, paper 8, 71–82.
1998a Conformal invariance of Voronoi percolation, *Communications in Mathematical Physics* 197, 75–107.
1998b Exceptional planes of percolation, *Probability Theory and Related Fields* 111, 551–564.

Berg, J. van den
1981 Percolation theory on pairs of matching lattics, *Journal of Mathematical Physics* 22, 153–157.
1982 A note on percolation theory, *Journal of Physics A: Mathematical and General* 15, 605–610.
1983 Unpublished.
1985 Disjoint occurrences of events: results and conjectures, *Particle Systems, Random Media and Large Deviations* (R. T. Durrett, ed.), Contemporary Mathematics no. 41, AMS, Providence, R. I., pp. 357–361.
1996 A note on disjoint-occurrence inequalities for marked Poisson point processes, *Journal of Applied Probability* 33, 420–426.

Berg, J. van den, Ermakov, A.
1996 A new lower bound for the critical probability of site percolation on the square lattice, *Random Structures and Algorithms* 8, 199–212.

Berg, J. van den, Fiebig, U.
1987 On a combinatorial conjecture concerning disjoint occurrences of events, *Annals of Probability* 15, 354–374.

Berg, J. van den, Grimmett, G. R., Schinazi, R. B.
1998 Dependent random graphs and spatial epidemics, *Annals of Applied Probability* 8, 317–336.

Berg, J. van den, Keane, M.
1984 On the continuity of the percolation probability function, *Particle Systems, Random Media and Large Deviations* (R. T. Durrett, ed.), Contemporary Mathematics Series, vol. 26, AMS, Providence, R. I., pp. 61–65.

Berg, J. van den, Kesten, H.
1985 Inequalities with applications to percolation and reliability, *Journal of Applied Probability* 22, 556–569.

1993 Inequalities for the time constant in first-passage percolation, *Advances in Applied Probability* 3, 56–80.

Berg, J. van den, Maes, C.
1994 Disagreement percolation in the study of Markov fields, *Annals of Probability* 22, 749–763.

Berg, J. van den, Steif, J.
1994 Percolation and the hard-core lattice gas model, *Stochastic Processes and their Applications* 49, 179–197.

Berlyand, L., Wehr, J.
1995 The probability distribution of the percolation threshold in a large system, *Journal of Physics A: Mathematical and General* 28, 7127–7133.
1997 Non-Gaussian limiting behavior of the percolation threshold in a large system, *Communications in Mathematical Physics* 185, 73–92.

Bezuidenhout, C. E., Grimmett, G. R.
1990 The critical contact process dies out, *Annals of Probability* 18, 1462–1482.
1991 Exponential decay for subcritical contact and percolation processes, *Annals of Probability* 19, 984–1009.
1997 A central limit theorem for random walks in random labyrinths (to appear).

Bezuidenhout, C. E., Grimmett, G. R., Kesten, H.
1993 Strict inequality for critical values of Potts models and random-cluster processes, *Communications in Mathematical Physics* 158, 1–16.

Bezuidenhout, C. E., Grimmett, G. R., Löffler, A.
1998 Percolation and minimal spanning trees, *Journal of Statistical Physics* 92, 1–34.

Biggs, N. L.
1976 *Interaction Models*, London Mathematical Society Lecture Notes no. 30, Cambridge University Press, Cambridge.

Bollobás, B., Grimmett, G. R., Janson, S.
1996 The random-cluster process on the complete graph, *Probability Theory and Related Fields* 104, 283–317.

Bolthausen, E.
1982 On the central limit theorem for stationary mixing random fields, *Annals of Probability* 10, 1047–1050.

Bondy, J. A., Murty, U. S. R.
1976 *Graph Theory with Applications*, Macmillan, New York.

Borgs, C., Chayes, J. T.
1996 The covariance matrix of the Potts model: a random cluster analysis, *Journal of Statistical Physics* 82, 1235–1297.

Borgs, C., Chayes, J. T., Kesten, H., Spencer, J.
1997 Uniform boundedness of critical crossing probabilities implies hyperscaling (to appear).
1998 The birth of the infinite cluster: finite-size scaling in percolation (to appear).

Bramson, M., Durrett, R. T., Schonmann, R. H.
1991 The contact process in a random environment, *Annals of Probability* 19, 960–983.

Brånvall G.
1980 A note on limit theorems in percolation, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 53, 317–328.

Broadbent, S. R.
1954 In discussion of Symposium on Monte Carlo Methods, *Journal of the Royal Statistical Society* (B) 16, 68.

Broadbent, S. R., Hammersley, J. M.
1957 Percolation processes I. Crystals and mazes, *Proceedings of the Cambridge Philosophical Society* 53, 629–641.

Bunimovitch, L. A., Sinai, Ya. G.
1981 Statistical properties of Lorentz gas with periodic configuration of scatterers, *Communications in Mathematical Physics* 78, 479–497.

Bunimovitch, L. A., Troubetzkoy, S. E.
1992 Recurrence properties of Lorentz lattice gas cellular automata, *Journal of Statistical Physics* 67, 289–302.
1994 Rotators, periodicity and absence of diffusion in cyclic cellular automata, *Journal of Statistical Physics* 74, 1–10.

Burton, R. M., Keane, M.
1989 Density and uniqueness in percolation, *Communications in Mathematical Physics* 121, 501–505.

Burton, R. M., Meester, R.
1993 Long range percolation in stationary point processes, *Random Structures and Algorithms* 4, 177–190.

Burton, R. M., Pemantle, R.
1993 Local characteristics, entropy and limit theorems for uniform spanning trees and domino tilings via transfer-impedances, *Annals of Probability* 21, 1329–1371.

Campanino, M., Chayes, J. T., Chayes, L.
1991 Gaussian fluctuations of connectivities in the subcritical regime of percolation, *Probability Theory and Related Fields* 88, 269–341.

Campanino, M., Russo, L.
1985 An upper bound for the critical probability for the three-dimensional cubic lattice, *Annals of Probability* 13, 478–491.

Cardy, J.
1992 Critical percolation in finite geometries, *Journal of Physics A: Mathematical and General* 25, L201.

Cerf, R.
1998a Large deviations of the finite cluster shape for two-dimensional percolation in the Hausdorff and L^1 topology (to appear).
1998b Large deviations for three dimensional supercritical percolation (to appear).

Chandler, R., Koplick, J., Lerman, K., Willemsen, J. F.
1982 Capillary displacement and percolation in porous media, *Journal of Fluid Mechanics* 119, 249–267.

Chayes, J. T., Chayes, L.
1986a Percolation and random media, *Critical Phenomena, Random Systems and Gauge Theories* (K. Osterwalder and R. Stora, eds.), Les Houches, Session XLIII, 1984, Elsevier, Amsterdam, pp. 1001–1142.
1986b Critical points and intermediate phases on wedges of $\mathbb{Z}^d$, *Journal of Physics A: Mathematical and General* 19, 3033–3048.
1986c An inequality for the infinite cluster density in percolation, *Physical Review Letters* 56, 1619–1622.

1986d Bulk transport properties and exponent inequalities for random resistor and flow networks, *Communications in Mathematical Physics* 105, 133–152.
1987a The mean field bound for the order parameter of Bernoulli percolation, *Percolation Theory and Ergodic Theory of Infinite Particle Systems* (H. Kesten, ed.), IMA Volumes in *Mathematics and its Applications*, vol. 8, Springer, New York, pp. 49–71.
1987b On the upper critical dimension of Bernoulli percolation, *Communications in Mathematical Physics* 113, 27–48.
1989 The large-N limit of the threshold values in Mandelbrot's fractal percolation process, *Journal of Physics A: Mathematical and General* 22, L501–L506.

Chayes, J. T., Chayes, L., Durrett, R. T.
1986 Critical behavior of the two-dimensional first passage time, *Journal of Statistical Physics* 45, 933–948.
1987 Inhomogeneous percolation problems and incipient infinite clusters, *Journal of Physics A: Mathematical and General* 20, 1521–1530.
1988 Connectivity properties of Mandelbrot's percolation process, *Probability Theory and Related Fields* 77, 307–324.

Chayes, J. T., Chayes, L., Fisher, D. S., Spencer, T.
1986 Finite-size scaling and correlation lengths for disordered systems, *Physical Review Letters* 57, 2999–3002.
1989 Correlation length bounds for disordered Ising ferromagnets, *Communications in Mathematical Physics* 120, 501–523.

Chayes, J. T., Chayes, L., Fröhlich, J.
1985 The low-temperature behavior of disordered magnets, *Communications in Mathematical Physics* 100, 399–437.

Chayes, J. T., Chayes, L., Grannan, E., Swindle, G.
1991 Phase transitions in Mandelbrot's percolation process in 3 dimensions, *Probability Theory and Related Fields* 90, 291–300.

Chayes, J. T., Chayes, L., Grimmett, G. R., Kesten, H., Schonmann, R. H.
1989 The correlation length for the high density phase of Bernoulli percolation, *Annals of Probability* 17, 1277–1302.

Chayes, J. T., Chayes, L., Kotecký, R.
1995 The analysis of the Widom–Rowlinson model by stochastic geometric methods, *Communications in Mathematical Physics* 172, 551–569.

Chayes, J. T., Chayes, L., Newman, C. M.
1984 Unpublished.
1985 The stochastic geometry of invasion percolation, *Communications in Mathematical Physics* 101, 383–407.
1987 Bernoulli percolation above threshold: an invasion percolation analysis, *Annals of Probability* 15, 1272–1287.

Chayes, J. T., Chayes, L., Schonmann, R. H.
1987 Exponential decay of connectivities in the two dimensional Ising model, *Journal of Statistical Physics* 49, 433–445.

Chayes, L.
1991 On the critical behavior of the first passage time in $d = 3$, *Helvetica Physica Acta* 64, 1055–1069.
1995a On the absence of directed fractal percolation, *Journal of Physics A: Mathematical and General* 28, L295–L301.

1995b Aspects of the fractal percolation process, *Fractal Geometry and Stochastics* (C. Bandt, S. Graf and M. Zähle, eds.), Birkhäuser, Boston, pp. 113–143.

1996 On the length of the shortest crossing in the super-critical phase of Mandelbrot's percolation process, *Stochastic Processes and their Applications* 61, 25–43.

Chow, Y. S., Teicher, H.

1978 *Probability Theory*, Springer, Berlin.

Cohen, E. G. D.

1991 New types of diffusions in lattice gas cellular automata, *Microscopic Simulations of Complex Hydrodynamical Phenomena* (M. Mareschal and B. L. Holian, eds.), Plenum Press, New York, pp. 137–152.

Cohen, E. G. D., Wang, F.

1995a New results for diffusion in Lorentz lattice gas cellular automata, *Journal of Statistical Physics* 81, 445–466.

1995b Novel phenomena in Lorentz lattice gases, *Physica A* 219, 56–87.

Coniglio, A.

1985 Shapes, surfaces and interfaces in percolation clusters, *Physics of Finely Divided Matter* (N. Boccara and M. Daoud, eds.), Les Houches, Session XLIV, 1985, Springer, Berlin, pp. 84–101.

Cox, J. T., Durrett, R.

1981 Some limit theorems for percolation processes with necessary and sufficient conditions, *Annals of Probability* 9, 583–603.

1983 Oriented percolation in dimensions $d \geq 4$: bounds and asymptotic formulas, *Mathematical Proceedings of the Cambridge Philosophical Society* 93, 151–162.

1988 Limit theorems for the spread of epidemics and forest fires, *Stochastic Processes and their Applications* 30, 171–191.

Cox, J. T., Grimmett, G. R.

1981 Central limit theorems for percolation models, *Journal of Statistical Physics* 25, 237–251.

1984 Central limit theorems for associated random variables and the percolation model, *Annals of Probability* 12, 514–528.

Dal Maso, G., Modica, L.

1986 Nonlinear stochastic homogenization and ergodic theory, *Journal für die Reine und Angewandte Mathematik* 368, 28–42.

Dekking, F. M., Grimmett, G. R.

1988 Superbranching processes and projections of random Cantor sets, *Probability Theory and Related Fields* 78, 335–355.

Dekking, F. M., Meester, R. W. J.

1990 On the structure of Mandelbrot's percolation process and other random Cantor sets, *Journal of Statistical Physics* 58, 1109–1126.

Delyon, F.

1980 Taille, forme et nombre des amas dans les problèmes de percolation, Thèse de 3me cycle, Université Pierre et Marie Curie, Paris.

DeMasi, A., Ferrari, P. A., Goldstein, S., Wick, W. D.

1985 Invariance principle for reversible Markov processes with application to diffusion in the percolation regime, *Particle Systems, Random Media and Large Deviations* (R. T. Durrett, ed.), Contemporary Mathematics no. 41, AMS, Providence, R. I., pp. 71–85.

1989 An invariance principle for reversible Markov processes. Applications to random motions in random environments, *Journal of Statistical Physics* 55, 787–855.

Deuschel, J.-D., Pisztora, A.

1996 Surface order large deviations for high-density percolation, *Probability Theory and Related Fields* 104, 467–482.

Deutscher, G., Zallen, R., Adler, J.

1983 *Percolation Structures and Processes*, editors, Annals of the Israel Physical Society 5.

Diao, Y., Janse van Rensburg, E. J.

1998 Percolation of linked circles, *Topology and Geometry in Polymer Science*, IMA Volumes in *Mathematics and its Applications*, vol. 103, Springer, New York, pp. 79–88.

Dunford, N., Schwartz, J. T.

1958 *Linear Operators*, Volume 1, Wiley–Interscience, New York.

Durrett, R. T.

1984 Oriented percolation in two dimensions, *Annals of Probability* 12, 999–1040.

1985a Some general results concerning the critical exponents of percolation processes, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 69, 421–437.

1985b *Particle Systems, Random Media and Large Deviations*, editor, Contemporary Mathematics 41, AMS, Providence, R. I..

1988 *Lecture Notes on Particle Systems and Percolation*, Wadsworth and Brooks/Cole, Pacific Grove, California.

1991 The contact process, 1974–1989, *Mathematics of Random Media,* Blacksburg, VA, Lectures in Applied Mathematics, vol. 27, AMS, Providence, R. I., pp. 1–18.

1995 Ten lectures on particle systems, *Ecole d'Eté de Probabilités de Saint Flour* XXIII-1993 (P. Bernard, ed.), Lecture Notes in Mathematics no. 1608, Springer, Berlin, pp. 97–201.

Durrett, R. T., Kesten, H.

1990 The critical parameter for connectedness of some random graphs, *A Tribute to Paul Erdős* (A. Baker, B. Bollobás, A. Hajnal, eds.), Cambridge University Press, Cambridge, pp. 161–176.

Durrett, R. T., Liggett, T. M.

1981 The shape of the limit set in Richardson's growth process, *Annals of Probability* 9, 186–193.

Durrett, R., Nguyen, B.

1985 Thermodynamic inequalities for percolation, *Communications in Mathematical Physics* 99, 253–269.

Durrett, R. T., Schonmann, R. H.

1987 Stochastic growth models, *Percolation Theory and Ergodic Theory of Infinite Particle Systems* (H. Kesten, ed.), IMA Volumes in *Mathematics and its Applications*, vol. 8, Springer, New York, pp. 85–119.

1988a Large deviations for the contact process and two dimensional percolation, *Probability Theory and Related Fields* 77, 583–603.

1988b The contact process on a finite set II, *Annals of Probability* 16, 1570–1583.

Durrett, R. T., Schonmann, R. H., Tanaka, N. I.

1989 Correlation lengths for oriented percolation, *Journal of Statistical Physics* 55, 965–979.

Dyson, F. J.
1969a Existence of a phase-transition in a one-dimensional Ising ferromagnet, *Communications in Mathematical Physics* 12, 91–107.
1969b Non-existence of spontaneous magnetization in a one-dimensional Ising ferromagnet, *Communications in Mathematical Physics* 12, 212–215.
1971 Phase transitions in ferromagnets, *Statistical Mechanics at the Turn of the Decade* (E. G. D. Cohen, ed.), Dekker, New York, pp. 129–144.

Edwards, R. G., Sokal, A. D.
1988 Generalization of the Fortuin–Kasteleyn–Swendsen–Wang representation and Monte Carlo algorithm, *The Physical Review* D 38, 2009–2012.

Efros, A. L.
1986 *Physics and Geometry of Disorder*, Mir, Moscow.

Ehrenfest, P.
1959 *Collected Scientific Papers* (M. J. Klein, ed.), North-Holland, Amsterdam.

Esary, J. D., Proschan, F.
1963 Coherent structures of non-identical components, *Technometrics* 5, 191–209.

Essam, J. W.
1971 Graph theory and statistical physics, *Discrete Mathematics* 1, 83–112.
1972 Percolation and cluster size, *Phase Transitions and Critical Phenomena* (C. Domb and M. S. Green, eds.), Volume II, Academic Press, London, pp. 197–270.
1980 Percolation theory, *Reports on Progress in Physics* 43, 833–949.

Falconer, K. J.
1986 Random fractals, *Mathematical Proceedings of the Cambridge Philosophical Society* 100, 559–582.

Falconer, K. J., Grimmett, G. R.
1992 On the geometry of random Cantor sets and fractal percolation, *Journal of Theoretical Probability* 5, 465–485; Erratum (1994) 7, 209–210.

Feller, W.
1968 *An Introduction to Probability Theory*, Volume I, 3rd Edition, Wiley, New York.
1971 *An Introduction to Probability Theory*, Volume II, 2nd Edition, Wiley, New York.

Fill, J., Pemantle, R.
1992 Oriented percolation, first-passage percolation and covering times for Richardson's model on the n-cube, *Annals of Applied Probability* 3, 593–629.

Fisher, M. E.
1961 Critical probabilities for cluster size and percolation problems, *Journal of Mathematical Physics* 2, 620–627.
1983 Scaling, universality and renormalization group, *Critical Phenomena* (F. J. W. Hahne, ed.), Lecture Notes in Physics no. 186, Springer, Berlin, pp. 1–139.

Fisher, M. E., Essam, J. W.
1961 Some cluster size and percolation problems, *Journal of Mathematical Physics* 2, 609–619.

Flory, P. J.
1941 Molecular size distribution in three dimensional gelation I–III, *Journal of the American Chemical Society* 63, 3083–3100.

Fontes, L., Newman, C. M.
1993 First passage percolation for random colorings of $\mathbb{Z}^d$, *Annals of Applied Probability* 3, 746–762; Erratum 4, 254.

Ford, L. R., Fulkerson, D. R.
1962 *Flow in Networks*, Princeton University Press, Princeton.

Fortuin, C. M.
1971 *On the random-cluster model*, Doctoral thesis, University of Leiden.
1972a On the random cluster model. II. The percolation model, *Physica* 58, 393–418.
1972b On the random cluster model. III. The simple random-cluster process, *Physica* 59, 545–570.

Fortuin, C. M., Kasteleyn, P. W.
1972 On the random cluster model. I. Introduction and relation to other models, *Physica* 57, 536–564.

Fortuin, C. M., Kasteleyn, P. W., Ginibre, J.
1971 Correlation inequalities on some partially ordered sets, *Communications in Mathematical Physics* 22, 89–103.

Friedgut, E., Kalai, G.
1996 Every monotone graph property has a sharp threshold, *Proceedings of the American Mathematical Society* 124, 2993–3002.

Frisch, H. L., Gordon, S. B., Vyssotsky, V. A., Hammersley, J. M.
1962 Monte Carlo solution of bond percolation processes on various crystal lattices, *The Bell System Technical Journal* 41, 909–920.

Frisch, H. L., Hammersley, J. M.
1963 Percolation processes and related topics, *Journal of the Society for Industrial and Applied Mathematics* 11, 894–918.

Fröhlich, J., Spencer, T.
1982 The phase transition in the one-dimensional Ising model with $1/r^2$ interaction energy, *Communications in Mathematical Physics* 84, 87–101.

Gandolfi, A.
1989 Uniqueness of the infinite cluster for stationary finite range Gibbs states, *Annals of Probability* 17, 1403–1415.

Gandolfi, A., Grimmett, G. R., Russo, L.
1988 On the uniqueness of the infinite open cluster in the percolation model, *Communications in Mathematical Physics* 114, 549–552.

Gandolfi, A., Keane, M., Newman, C. M.
1992 Uniqueness of the infinite component in a random graph with applications to percolation and spin glasses, *Probability Theory and Related Fields* 92, 511–527.

Gandolfi, A., Keane, M., Russo, L.
1988 On the uniqueness of the infinite occupied cluster in dependent two-dimensional site percolation, *Annals of Probability* 16, 1147–1157.

Gandolfi, A., Kesten, H.
1994 Greedy lattice animals II: Linear growth, *Advances in Applied Probability* 4, 76–107.

Gennes, P. G. de
1976 La percolation: un concept unificateur, *La Recherche* 7, 919–927.

Gennes, P. G. de, Guyon, E.
1978 Lois générales pour l'injection d'un fluide dans un milieu poreux aléatoire, *Journal de Mécanique* 119, 249–267.

Gilbert, E. N.
1961 Random plane networks, *Journal of the Society for Industrial and Applied Mathematics* 9, 533–543.

Gluck, H.
1974 Almost all simply connected closed surfaces are rigid, *Geometric Topology*, Lecture Notes in Mathematics no. 438, Springer, Berlin, pp. 225–239.

Godsil, C., McKay, B.
1980 The dimension of a graph, *Quarterly Journal of Mathematics, Oxford* 31, 423–427.

Golden, K., Papanicolaou, G.
1982 Bounds for effective parameters of heterogeneous media by analytic continuation, *Communications in Mathematical Physics* 90, 473–491.

Gordon, D. M.
1991 Percolation in high dimensions, *Bulletin of the London Mathematical Society* 44, 373–384.

Graf, S.
1987 Statistically self-similar fractals, *Probability Theory and Related Fields* 74, 357–392.

Graf, S., Mauldin, R. D., Williams, S. C.
1987 Exact Hausdorff dimension in random recursive constructions, *Memoirs of the American Mathematical Society* 381.

Graver, J., Servatius, B., Servatius, H.
1993 *Combinatorial Rigidity*, Graduate Studies in Mathematics no. 2, AMS, Providence, R. I..

Greene, J. W., El Gamal, A.
1984 Configuration of VLSI arrays in the presence of defects, *Journal of the Association for Computing Machinery* 31, 694–717.

Griffiths, R. B.
1967 Correlations in Ising ferromagnets. II. External magnetic fields, *Journal of Mathematical Physics* 8, 484–489.

Grigorchuk, R. I.
1983 On Milnor's problem of group growth, *Soviet Mathematics Doklady* 28, 23–26.

Grimmett, G. R.
1976 On the number of clusters in the percolation model, *Journal of the London Mathematical Society* 13, 346–350.
1978a The rank functions of large random lattices, *Journal of the London Mathematical Society* 18, 567–575.
1978b Multidimensional lattices and their partition functions, *Quarterly Journal of Mathematics, Oxford* 29, 142–157.
1979 Central limit theorems in percolation theory, unpublished manuscript.
1981a Critical sponge dimensions in percolation theory, *Advances in Applied Probability* 13, 314–324.
1981b On the differentiability of the number of clusters per site in the percolation model, *Journal of the London Mathematical Society* 23, 372–384.
1981c Unpublished.
1983 Bond percolation on subsets of the square lattice, the transition between one-dimensional and two-dimensional behaviour, *Journal of Physics A: Mathematical and General* 16, 599–604.

1985a On a conjecture of Hammersley and Whittington concerning bond percolation on subsets of the simple cubic lattice, *Journal of Physics A: Mathematical and General* 18, L49–L52.
1985b The largest components in a random lattice, *Studia Scientiarum Mathematicarum Hungarica* 20, 325–331.
1985c Random flows: network flows and electrical flows through random media, *Surveys in Combinatorics* (I. Anderson, ed.), Cambridge University Press, Cambridge, pp. 59–96.
1986 Percolation processes and dimensionality, *Stochastic Spatial Processes* (P. Tautu, ed.), Lecture Notes in Mathematics no. 1212, Springer, Berlin, pp. 165–174.
1987a Long paths and cycles in a random lattice, *Random Graphs '85*, Annals of Discrete Mathematics 33, North-Holland, Amsterdam, pp. 69–76.
1987b Recent progress in percolation, *Proceedings of the First World Congress of the Bernoulli Society* (Yu. V. Prohorov and V. V. Sazonov, eds.), Tashkent 1986, Volume 1, VNU Science Press, Utrecht, pp. 353–358.
1989 *Percolation*, 1st Edition, Springer, Berlin.
1993 Differential inequalities for Potts and random-cluster processes, *Cellular Automata and Cooperative Systems* (N. Boccara et al., eds.), Kluwer, Dordrecht, pp. 227–236.
1994a Potts models and random-cluster processes with many-body interactions, *Journal of Statistical Physics* 75, 67–121.
1994b The random-cluster model, *Probability, Statistics and Optimisation* (F. P. Kelly, ed.), Wiley, Chichester, pp. 49–63.
1994c Percolative problems, *Probability and Phase Transition* (G. R. Grimmett, ed.), Kluwer, Dordrecht, pp. 69–86.
1994d *Probability and Phase Transition*, editor, Kluwer, Dordrecht.
1995a Comparison and disjoint-occurrence inequalities for random-cluster models, *Journal of Statistical Physics* 78, 1311–1324.
1995b The stochastic random-cluster process and the uniqueness of random-cluster measures, *Annals of Probability* 23, 1461–1510.
1997 Percolation and disordered systems, *Ecole d'Eté de Probabilités de Saint Flour XXVI–1996* (P. Bernard, ed.), Lecture Notes in Mathematics no. 1665, Springer, Berlin, pp. 153–300.
1999a Percolation, *Development of Mathematics 1950–2000*, Birkhäuser (to appear).
1999b Stochastic pin-ball, *Proceedings of the Cortona Workshop on Random Walks and Discrete Potential Theory*, Cambridge University Press, Cambridge (to appear).
1999c Inequalities and entanglements for percolation and random-cluster models (to appear).

Grimmett, G. R., Holroyd, A. E.
1998 Entanglement in percolation (to appear).

Grimmett, G. R., Keane, M., Marstrand, J. M.
1984 On the connectedness of a random graph, *Mathematical Proceedings of the Cambridge Philosophical Society* 96, 151–166.

Grimmett, G. R., Kesten, H.
1984a First-passage percolation, network flows and electrical resistances, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 66, 335–366.
1984b Random electrical networks on complete graphs, *Journal of the London Mathematical Society* 30, 171–192.

Grimmett, G. R., Kesten, H., Zhang, Y.
1993 Random walk on the infinite cluster of the percolation model, *Probability Theory and Related Fields* 96, 33–44.

Grimmett, G. R., Marstrand, J. M.
1990 The supercritical phase of percolation is well behaved, *Proceedings of the Royal Society (London), Series A* 430, 439–457.

Grimmett, G. R., Menshikov, M. V., Volkov, S. E.
1996 Random walks in random labyrinths, *Markov Processes and Related Fields* 2, 69–86.

Grimmett, G. R., Newman, C. M.
1990 Percolation in $\infty + 1$ dimensions, *Disorder in Physical Systems* (G. R. Grimmett and D. J. A. Welsh, eds.), Clarendon Press, Oxford, pp. 167–190.

Grimmett, G. R., Piza, M. S. T.
1997 Decay of correlations in subcritical Potts and random-cluster models, *Communications in Mathematical Physics* 189, 465–480.

Grimmett, G. R., Stacey, A. M.
1998 Critical probabilities for site and bond percolation models, *Annals of Probability* (to appear).

Grimmett, G. R., Stirzaker, D. R.
1992 *Probability and Random Processes*, 2nd Edition, Oxford University Press, Oxford.

Häggström, O.
1996 The random-cluster model on a homogeneous tree, *Probability Theory and Related Fields* 104, 231–253.
1997 Infinite clusters in dependent automorphism invariant percolation on trees, *Annals of Probability* 25, 1423–1436.
1998 Random-cluster representations in the study of phase transitions, *Markov Processes and Related Fields* (to appear).

Häggström, O., Meester, R.
1995 Asymptotic shapes in stationary first passage percolation, *Annals of Probability* 23, 1511–1522.
1996 Nearest neighbour and hard sphere models in continuum percolation, *Random Structures and Algorithms* 9, 295–315.

Häggström, O., Mossel, E.
1998 Nearest neighbor walks with low predictability profile and percolation in $2 + \varepsilon$ dimensions, *Annals of Probability* 26, 1212–1231.

Häggström, O., Pemantle, R.
1998 First-passage percolation and a model for competing spatial growth, *Journal of Applied Probability* 35, 683–692.

Häggström, O., Peres, Y.
1998 Monotonicity of uniqueness for percolation on Cayley graphs: All infinite clusters are born simultaneously, *Probability Theory and Related Fields* (to appear).

Häggström, O., Peres, Y., Steif, J.
1997 Dynamical percolation, *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* 33, 497–528.

Hall, P.
1985 On continuum percolation, *Annals of Probability* 13, 1250–1266.
1986 Clump counts in a mosaic, *Annals of Probability* 14, 424–458.
1988 *An Introduction to the Theory of Coverage Processes*, Wiley, New York.

Halley, J. W.
1983 Polychromatic percolation, *Percolation Structures and Processes*, Annals of the Israel Physical Society 5, pp. 323–351.

Hammersley, J. M.
1957a Percolation processes. Lower bounds for the critical probability, *Annals of Mathematical Statistics* 28, 790–795.
1957b Percolation processes II. The connective constant, *Proceedings of the Cambridge Philosophical Society* 53, 642–645.
1959 Bornes supérieures de la probabilité critique dans un processus de filtration, *Le Calcul de Probabilités et ses Applications*, CNRS, Paris, pp. 17–37.
1961 Comparison of atom and bond percolation, *Journal of Mathematical Physics* 2, 728–733.
1962 Generalization of the fundamental theorem on subadditive functions, *Proceedings of the Cambridge Philosophical Society* 58, 235–238.
1963 A Monte Carlo solution of percolation in a cubic lattice, *Methods in Computational Physics* (B. Alder, S. Fernbach, M. Rotenberg, eds.), Volume I, Academic Press, London, pp. 281–298.
1980 A generalization of McDiarmid's theorem for mixed Bernoulli percolation, *Proceedings of the Cambridge Philosophical Society* 88, 167–170.
1983 Origins of percolation theory, *Annals of the Israel Physical Society* 5, 47–57.
1988 Mesoadditive processes and the specific conductivity of lattices, *Journal of Applied Probability* 25A, 347–358.

Hammersley, J. M., Morton, W.
1954 Poor man's Monte Carlo, *Journal of the Royal Statistical Society* (B) 16, 23–38.

Hammersley, J. M., Welsh, D. J. A.
1965 First-passage percolation, subadditive processes, stochastic networks and generalized renewal theory, *Bernoulli, Bayes, Laplace Anniversary Volume* (J. Neyman and L. LeCam, eds.), Springer, Berlin, pp. 61–110.
1980 Percolation theory and its ramifications, *Contemporary Physics* 56, 13–20.

Hammersley, J. M., Whittington, S. G.
1985 Self-avoiding walks in wedges, *Journal of Physics A: Mathematical and General* 18, 101–111.

Hankey, A.
1978 Three properties of the infinite cluster in percolation theory, *Journal of Physics A: Mathematical and General* 11, L49–L55.

Hara, T.
1990 Mean-field critical behaviour for correlation length for percolation in high dimensions, *Probability Theory and Related Fields* 86, 337–385.

Hara, T., Slade, G.
1989a Mean-field critical behaviour of percolation in high dimensions, *Proceedings of the IXth International Congress on Mathematical Physics* (B. Simon, A. Truman, I. M. Davies, eds.), Adam Hilger, Bristol, pp. 450–453.
1989b The triangle condition for percolation, *Bulletin of the American Mathematical Society* 21, 269–273.

1990 Mean-field critical behaviour for percolation in high dimensions, *Communications in Mathematical Physics* 128, 333–391.
1994 Mean-field behaviour and the lace expansion, *Probability and Phase Transition* (G. R. Grimmett, ed.), Kluwer, Dordrecht, pp. 87–122.
1995 The self-avoiding walk and percolation critical points in high dimensions, *Combinatorics, Probability, Computing* 4, 197–215.
1998 The incipient infinite cluster in high-dimensional percolation (to appear).

Harris, T. E.
1960 A lower bound for the critical probability in a certain percolation process, *Proceedings of the Cambridge Philosophical Society* 56, 13–20.
1963 *The Theory of Branching Processes*, Springer, Berlin.
1974 Contact interactions on a lattice, *Annals of Probability* 2, 969–988.
1978 Additive set-valued Markov processes and graphical methods, *Annals of Probability* 6, 355–378.

Hauge, E. H., Cohen, E. G. D.
1967 Normal and abnormal effects in Ehrenfest wind–tree model, *Physics Letters* 25A, 78–79.

Hawkes, J.
1981 Trees generated by a simple branching process, *Journal of the London Mathematical Society* 24, 373–384.

Herrndorf, N.
1985 Mixing conditions for random fields in percolation theory, unpublished manuscript.

Higuchi, Y.
1982 Coexistence of the infinite (∗) clusters: a remark on the square lattice site percolation, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 61, 75–81.
1989 A remark on the percolation for the 2D Ising model, *Osaka Journal of Mathematics* 26, 207–224.
1992 *Percolation* (in Japanese), Yuseisha/Seiunsha.
1993a A sharp transition for the two-dimensional Ising percolation, *Probability Theory and Related Fields* 97, 489–514.
1993b Coexistence of infinite *-clusters II: Ising percolation in two dimensions, *Probability Theory and Related Fields* 97, 1–34.
1995 Ising model percolation, *Sugaku Expositions* 47, 111–124.
1996 Mixing condition and percolation for Ising models, *Probability Theory and Mathematical Statistics (Tokyo 1995)*, World Scientific, pp. 110–121.

Hille, E.
1948 *Functional Analysis and Semi-Groups*, AMS Colloquium Publications no. 31, New York.

Hintermann, D., Kunz, H., Wu, F. Y.
1978 Exact results for the Potts model in two dimensions, *Journal of Statistical Physics* 19, 623–632.

Hoffman, C., Mossel, E.
1998 Energy of flows on percolation clusters (to appear).

Hollander, W. Th. F. den, Keane, M.
1986 Inequalities of FKG type, *Physica* 138A, 167–182.

Holley, R.
1974 Remarks on the FKG inequalities, *Communications in Mathematical Physics* 36, 227–231.

Holmes, R. A.
1975 Random graphs with the natural numbers as vertices, unpublished manuscript.

Holroyd, A. E.
1998a Existence and uniqueness of infinite components in generic rigidity percolation, *Annals of Applied Probability* 8, 944–973.
1998b Existence of a phase transition for entanglement percolation (to appear).

Hughes, B. D.
1996 *Random Walks and Random Environments. Vol. 2. Random Environments*, Oxford University Press, New York.

Hughes, B. D., Ninham, B. W.
1983 *The Mathematics and Physics of Disordered Media*, editors, Lecture Notes in Mathematics no. 1035, Springer, Berlin.

Imbrie, J. Z., Newman, C. M.
1988 An intermediate phase with slow decay of correlations in one dimensional $1/|x-y|^2$ percolation, Ising and Potts models, *Communications in Mathematical Physics* 118, 303–336.

Ising, E.
1925 Beitrag zur Theorie des Ferromagnetismus, *Zeitschrift für Physik* 31, 253–258.

Jacobs, D. J., Thorpe, M. F.
1995 Generic rigidity percolation: the pebble game, *Physical Review Letters* 75, 4051–4054.
1996 Generic rigidity percolation in two dimensions, *Physical Review E* 53, 3682–3693.

Jikov, V. V., Kozlov, S. M., Oleinik, O. A.
1994 *Homogenization of Differential Operators and Integral Functionals*, Springer, Berlin.

Kalikow, S., Weiss, B.
1988 When are random graphs connected?, *Israel Journal of Mathematics* 62, 257–268.

Kantor, T., Hassold, G. N.
1988 Topological entanglements in the percolation problem, *Physical Review Letters* 60, 1457–1460.

Kapitulnik, A., Aharony, A., Deutscher, G., Stauffer, D.
1983 Self similarity and correlations in percolation, *Journal of Physics A: Mathematical and General* 16, L269–L274.

Kasteleyn, P. W., Fortuin, C. M.
1969 Phase transitions in lattice systems with random local properties, *Journal of the Physical Society of Japan* 26, 11–14, Supplement.

Kesten, H.
1980a The critical probability of bond percolation on the square lattice equals $\frac{1}{2}$, *Communications in Mathematical Physics* 74, 41–59.
1980b On the time constant and path length of first-passage percolation, *Advances in Applied Probability* 12, 848–863.
1980c Unpublished.

1981 Analyticity properties and power law estimates in percolation theory, *Journal of Statistical Physics* 25, 717–756.
1982 *Percolation Theory for Mathematicians*, Birkhäuser, Boston.
1985 First-passage percolation and a higher dimensional generalization, *Particle Systems, Random Media and Large Deviations* (R. Durrett, ed.), Contemporary Mathematics 41, AMS, Providence, R. I., pp. 235–251.
1986a Aspects of first-passage percolation, *Ecole d'Eté de Probabilités de Saint Flour* XIV-1984 (P. L. Hennequin, ed.), Lecture Notes in Mathematics no. 1180, Springer, Berlin, pp. 125–264.
1986b The incipient infinite cluster in two-dimensional percolation, *Probability Theory and Related Fields* 73, 369–394.
1986c Subdiffusive behavior of random walk on a random cluster, *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* 22, 425–487.
1987a Surfaces with minimal random weights and maximal flows: A higher dimensional version of first-passage percolation, *Illinois Journal of Mathematics* 31, 99–166.
1987b Scaling relations for 2D-percolation, *Communications in Mathematical Physics* 109, 109–156.
1987c A scaling relation at criticality for 2D-percolation, *Percolation Theory and Ergodic Theory of Infinite Particle Systems* (H. Kesten, ed.), IMA Volumes in *Mathematics and its Applications*, vol. 8, Springer, New York, pp. 203–212.
1987d *Percolation Theory and Ergodic Theory of Infinite Particle Systems*, editor, IMA Volumes in *Mathematics and its Applications*, vol. 8, Springer, New York.
1987e Percolation theory and first-passage percolation, *Annals of Probability* 15, 1231–1271.
1988a Recent results in rigorous percolation theory, *Astérisque* 157–158, 217–231.
1988b Correlation length and critical probabilities of slabs for percolation, unpublished manuscript.
1990 Asymptotics in high dimensions for percolation, *Disorder in Physical Systems* (G. R. Grimmett, D. J. A. Welsh, eds.), Clarendon Press, Oxford.
1991 Asymptotics in high dimensions for the Fortuin–Kasteleyn random cluster model, *Spatial Stochastic Processes* (K. S. Alexander and J. C. Watkins, eds.), Birkhäuser, Boston, pp. 57–85.
1992 Connectivity of certain graphs on halfspaces, quarter spaces, . . . , *Probability Theory* (L. H. Y. Chen, K. P. Choi, K. Hu, J.-H. Lou, eds.), de Gruyter, Berlin, pp. 91–104.
1993 On the speed of convergence in first-passage percolation, *Annals of Applied Probability* 3, 296–338.

Kesten, H., Schonmann, R. H.
1990 Behavior in large dimensions of the Potts and Heisenberg models, *Reviews in Mathematical Physics* 1, 147–182.

Kesten, H., Su, Z.
1998 Asymptotic behavior of the critical probability for ρ-percolation in high dimensions (to appear).

Kesten, H., Zhang, Y.
1987 Strict inequalities for some critical exponents in two-dimensional percolation, *Journal of Statistical Physics* 46, 1031–1055.
1990 The probability of a large finite cluster in supercritical Bernoulli percolation, *Annals of Probability* 18, 537–555.

1993 The tortuosity of occupied crossings of a box in critical percolation, *Journal of Statistical Physics* 70, 599–611.
1997 A central limit theorem for critical first passage percolation in two dimensions, *Probability Theory and Related Fields* 107, 137–160.

Kirkpatrick, S.
1979 Models of disordered materials, *La Matière Mal Condensée* (R. Balian et al., eds.), Les Houches, Session XXXI, 1978, North-Holland, Amsterdam.

Klein, A.
1994a Multiscale analysis in disordered systems: Percolation and contact process in a random environment, *Probability and Phase Transition* (G. R. Grimmett, ed.), Kluwer, Dordrecht, pp. 139–152.
1994b Extinction of contact and percolation processes in a random environment, *Annals of Probability* 22, 1227–1251.

Klein, S. T., Shamir, E.
1982 An algorithmic method for studying percolation clusters, *Stanford University Department of Computer Science Report* no. STAN-CS-82-933.

Kotecký, R., Shlosman, S.
1982 First order phase transitions in large entropy lattice systems, *Communications in Mathematical Physics* 83, 493–515.

Künnemann, R.
1983a Effective conductivity on a lattice as a limit of box conductivities, Ph. D. Thesis, University of Heidelberg.
1983b The diffusion limit for reversible jump processes on $\mathbb{Z}^d$ with ergodic random bond connectivities, *Communications in Mathematical Physics* 80, 27–68.

Kunz, H., Souillard, B.
1978 Essential singularity in percolation problems and asymptotic behavior of cluster size distribution, *Journal of Statistical Physics* 19, 77–106.

Kuulasmaa, K.
1982 The spatial general epidemic and locally dependent random graphs, *Journal of Applied Probability* 19, 745–758.

Laanait, L., Messager, A., Miracle-Sole, S., Ruiz, J., Shlosman, S.
1991 Interfaces in the Potts model I: Pirogov–Sinai theory of the Fortuin–Kasteleyn representation, *Communications in Mathematical Physics* 140, 81–91.

Laanait, L., Messager, A., Ruiz, J.
1986 Phase coexistence and surface tensions for the Potts model, *Communications in Mathematical Physics* 105, 527–545.

Lalley, S. P.
1998 Percolation on Fuchsian groups, *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* 34, 151–177.

Langlands, R. P.
1993 Dualität bei endlichen Modellen der Perkolation, *Mathematische Nachrichten* 160, 7–58.

Langlands, R. P., Lafortune, M.-A.
1994 Finite models for percolation, *Contemporary Mathematics* 177, 227–246.

Langlands, R., Pichet, C., Pouliot, P., Saint-Aubin, Y.
1992 On the universality of crossing probabilities in two-dimensional percolation, *Journal of Statistical Physics* 67, 553–574.

Langlands, R., Pouliot, P., Saint-Aubin, Y.
1994 Conformal invariance in two-dimensional percolation, *Bulletin of the American Mathematical Society* 30, 1–61.

Lenormand, R., Bories, S.
1980 Description d'un mécanisme de connexion de liaison destiné a l'étude du drainage avec piégeage en milieu poreux, *Comptes Rendus des Séances de l'Académie des Sciences* 291, *Série B*, 279–282.

Levin, D., Peres, Y.
1998 Energy and cutsets in infinite percolation clusters, *Proceedings of the Cortona Workshop on Random Walks and Discrete Potential Theory*, Cambridge University Press, Cambridge (to appear).

Licea, C., Newman, C. M.
1996 Geodesics in two-dimensional first-passage percolation, *Annals of Probability* 24, 399–410.

Licea, C., Newman, C. M., Piza, M. S. T.
1996 Superdiffusivity in first-passage percolation, *Probability Theory and Related Fields* 106, 559–591.

Lieb, E. H.
1980 A refinement of Simon's correlation inequality, *Communications in Mathematical Physics* 77, 127–135.

Liggett, T. M.
1985 *Interacting Particle Systems*, Springer, Berlin.
1991 Spatially inhomogeneous contact processes, *Spatial Stochastic Processes* (K. S. Alexander and J. C. Watkins, eds.), Birkhäuser, Boston, pp. 105–140.
1992 The survival of one-dimensional contact processes in random environments, *Annals of Probability* 20, 696–723.
1995a Improved upper bounds for the contact process critical value, *Annals of Probability* 23, 697–723.
1995b Survival of discrete time growth models, with applications to oriented percolation, *Annals of Applied Probability* 5, 613–636.
1996a Multiple transition points for the contact process on the binary tree, *Annals of Probability* 24, 1675–1710.
1996b Branching random walks and contact processes on trees, *Probability Theory and Related Fields* 106, 495–519.

Liggett, T. M., Schonmann, R. H., Stacey, A. M.
1997 Domination by product measures, *Annals of Probability* 25, 71–95.

Lorentz, H. A.
1905 The motion of electrons in metallic bodies, I, II, III, *Koninklijke Akademie van Wetenschappen te Amsterdam, Section of Sciences* 7, 438–453, 585–593, 684–691.

Łuczak, T., Wierman, J. C.
1988 Critical probability bounds for two-dimensional site percolation models, *Journal of Physics A: Mathematical and General* 21, 3131–3138.
1989 Counterexamples in AB percolation, *Journal of Physics A: Mathematical and General* 22, 185–191.

Lyons, R.
1990 Random walks and percolation on trees, *Annals of Probability* 18, 931–958.

1992 Random walks, capacity and percolation on trees, *Annals of Probability* 20, 2043–2088.

Lyons, R., Pemantle, R.
1992 Random walk in a random environment and first-passage percolation on trees, *Annals of Probability* 20, 125–137.

Madras, N., Schinazi, R. B., Schonmann, R. H.
1994 On the critical behavior of the contact process in deterministic inhomogeneous environments, *Annals of Probability* 22, 1140–1159.

Madras, N., Slade, G.
1993 *The Self-Avoiding Walk*, Birkhäuser, Boston.

Malyshev, V. A.
1975 The central limit theorem for Gibbsian random fields, *Soviet Mathematics Doklady* 16, 1141–1145.

Mandelbrot, B.
1983 *The Fractal Geometry of Nature*, W. H. Freeman, San Francisco.

Mauldin, R. D., Williams, S. C.
1986 Random recursive constructions: asymptotic geometry and topological properties, *Transactions of the American Mathematical Society* 295, 325–346.

McDiarmid, C. J. H.
1980 Clutter percolation and random graphs, *Mathematical Programming Study* 13, 17–25.
1981 General percolation and random graphs, *Advances in Applied Probability* 13, 40–60.
1983 General first-passage percolation, *Advances in Applied Probability* 15, 149–161.

McLeish, D. L.
1974 Dependent central limit theorems and invariance principles, *Annals of Probability* 2, 620–628.

Meester, R.
1989 An algorithm for calculating critical probabilities and percolation functions in percolation models defined by rotations, *Ergodic Theory and Dynamical Systems* 8, 495–509.
1992 Connectivity in fractal percolation, *Journal of Theoretical Probability* 5, 775–789.
1994 Uniqueness in percolation theory; a review, *Statistica Neerlandica* 48, 237–252.
1995 Equality of critical densities in continuum percolation, *Journal of Applied Probability* 32, 90–104.

Meester, R., Nowicki, T.
1989 Infinite clusters and critical values in two-dimensional circle percolation, *Israel Journal of Mathematics* 68, 63–81.

Meester, R., Penrose, M., Sarkar, A.
1997 The random connection model in high dimensions, *Statistics and Probability Letters* 35, 145–153.

Meester, R., Roy, R.
1994 Uniqueness of unbounded occupied and vacant components in Boolean models, *Advances in Applied Probability* 4, 933–951.
1996 *Continuum Percolation*, Cambridge University Press, Cambridge.

Meester, R., Roy, R., Sarkar, A.

1994 Non-universality and continuity of the critical covered volume fraction in continuum percolation, *Journal of Statistical Physics* 75, 123–134.

Meester, R., Steif, J.

1996 On the continuity of the critical value for long range percolation in the exponential case, *Communications in Mathematical Physics* 180, 483–504.

1998 Consistent estimation of percolation quantities, *Statistica Neerlandica* 52, 226–238.

Menshikov, M. V.

1985 Estimates of percolation thresholds for lattices in $\mathbb{R}^n$, *Soviet Mathematics Doklady* 32, 368–370.

1986 Coincidence of critical points in percolation problems, *Soviet Mathematics Doklady* 33, 856–859.

1987a Numerical bounds and strict inequalities for critical points of graphs and their subgraphs, *Theory of Probability and its Applications* (in Russian) 32, 599–602 (544–547 in translation).

1987b Critical points in the mathematical theory of percolation, Doctoral Dissertation, University of Moscow.

1987c Coincidence of critical points in percolation problems, *Soviet Mathematics Doklady* 33, 856–859.

1987d Quantitative estimates and rigorous inequalities for critical points of a graph and its subgraphs, *Theory of Probability and its Applications* 32, 544–547.

1987e Strict inequalities for critical probabilities of a regular graph and its subgraphs, *Uspekhi Mat. Nauk* 42, 239–240.

Menshikov, M. V., Molchanov, S. A., Sidorenko, A. F.

1986 Percolation theory and some applications, *Itogi Nauki i Techniki* (Series of Probability Theory, Mathematical Statistics, Theoretical Cybernetics) 24, 53–110.

Menshikov, M. V., Pelikh K. D.

1989 Percolation with several types of defects. Estimates of the critical probability on the square lattice, *Mathematical Notes* (in Russian) 46, 38–47 (778–785 in translation).

1993 Percolation method in a system of numerous particles, *Probabilistic Methods in Discrete Mathematics*, TVP Science Publishers, pp. 329–336.

Menshikov, M. V., Sidorenko, A. F.

1987 Coincidence of critical points for Poisson percolation models, *Theory of Probability and its Applications* (in Russian) 32, 603–606 (547–550 in translation).

Menshikov, M. V., Zuev, S. A.

1993 Models of ρ-percolation, *Probabilistic Methods in Discrete Mathematics* (V. F. Kolchin, ed.), TVP Science Publishers, pp. 337–347.

Moore, E. F., Shannon, C. E.

1956 Reliable circuits using less reliable components, *Journal of the Franklin Institute* 262, 191–208 and 281–297.

Neaderhouser, C. C.

1978 Limit theorems for multiply indexed mixing random variables with application to Gibbs random fields, *Annals of Probability* 6, 207–215.

1981 Some limit theorems for random fields, *Communications in Mathematical Physics* 61, 293–305.

Newman, C. M.
1980 Normal fluctuations and the FKG inequalities, *Communications in Mathematical Physics* 74, 119–128.
1986 Some critical exponent inequalities for percolation, *Journal of Statistical Physics* 45, 359–368.
1987a Percolation theory: a selective survey of rigorous results, *Advances in Multiphase Flow and Related Problems* (G. Papanicolaou, ed.), SIAM, Philadelphia, pp. 147–167.
1987b Inequalities for γ and related critical exponents in short and long range percolation, *Percolation Theory and Ergodic Theory of Infinite Particle Systems* (H. Kesten, ed.), IMA Volumes in *Mathematics and its Applications*, vol. 8, Springer, New York, pp. 229–244.
1987c Another critical exponent inequality for percolation: $\beta \geq 2/\delta$, *Journal of Statistical Physics* 47, 695–699.
1989 The $1/r^2$ phase transition: a review, *Proceedings of the IXth International Congress on Mathematical Physics* (B. Simon, A. Truman, I. M. Davies, eds.), Adam Hilger, Bristol, pp. 377–379.
1991a Ising models and dependent percolation, *Topics in Statistical Dependence* (H. W. Block, A. R. Sampson, T. H. Savits, eds.), IMS Lecture Notes, Monograph Series, vol. 16, pp. 395–401.
1991b Topics in percolation, *Mathematics of Random Media* (W. Kohler and B. White, eds.), vol. 27, AMS, Providence, R. I., pp. 41–51.
1994 Disordered Ising systems and random cluster representations, *Probability and Phase Transition* (G. R. Grimmett, ed.), Kluwer, Dordrecht, pp. 247–260.
1995 A surface view of first-passage percolation, *Proceedings of the International Congress of Mathematicians* (S. D. Chatterji, ed.), Birkhäuser, Boston, pp. 1017–1023.

Newman, C. M., Piza, M. S. T.
1995 Divergence of shape fluctuations in two dimensions, *Annals of Probability* 23, 977–1005.

Newman, C. M., Schulman, L. S.
1981a Infinite clusters in percolation models, *Journal of Statistical Physics* 26, 613–628.
1981b Number and density of percolating clusters, *Journal of Physics A: Mathematical and General* 14, 1735–1743.
1985 Infinite chains and clusters in one-dimensional directed and undirected percolation, *Particle Systems, Random Media and Large Deviations* (R. Durrett, ed.), Contemporary Mathematics 41, AMS, Providence, R. I., pp. 297–317.
1986 One-dimensional $1/|j-i|^s$ percolation models, *Communications in Mathematical Physics* 104, 547–571.

Newman, C. M., Wright, A. L.
1981 An invariance principle for certain dependent sequences, *Annals of Probability* 9, 671–675.

Newman, C. M., Wu, C. C.
1988 Unpublished.
1997 Percolation and contact processes with low-dimensional inhomogeneity, *Annals of Probability* 25, 1832–1845.

Nguyen, B. G.
1985 Correlation lengths for percolation processes, Ph. D. Thesis, University of California at Los Angeles.

1987a Correlation length and its critical exponent for percolation processes, *Journal of Statistical Physics* 46, 517–523.
1987b Gap exponents for percolation processes, *Journal of Statistical Physics* 49, 235–243.
1988 Fourier transform of the percolation free energy, *Probability Theory and Related Fields* 78, 165–168.

Nienhuis, B., Riedel, E. K., Schick, M.
1980 Magnetic exponents of the two-dimensional q-state Potts model, *Journal of Physics A: Mathematical and General* 13, L189–L192.

Nijs, M. P. M. den
1979 A relation between the temperature exponents of the eight-vertex and q-state Potts model, *Journal of Physics A: Mathematical and General* 12, 1857–1868.

Ornstein, L. S., Zernike, F.
1915 Accidental deviations of density and opalescence at the critical point of a single substance, *Koninklijke Akademie van Wetenschappen te Amsterdam, Section of Sciences* 17, 793–806.

Orzechowski, M. E.
1996 On the phase transition to sheet percolation in random Cantor sets, *Journal of Statistical Physics* 82, 1081–1098.
1998 A lower bound on the box dimension of crossings in fractal percolation, *Stochastic Processes and their Applications* 74, 53–65.

Pearson, R. P.
1980 Conjecture for the extended Potts model magnetic eigenvalue, *Physical Review B* 22, 2579–2580.

Peierls, R.
1936 On Ising's model of ferromagnetism, *Proceedings of the Cambridge Philosophical Society* 36, 477–481.

Pemantle, R.
1991 Choosing a spanning tree for the integer lattice uniformly, *Annals of Probability* 19, 1559–1574.
1992 The contact process on trees, *Annals of Probability* 20, 2089–2116.
1995 Uniform random spanning trees, *Topics in Contemporary Probability and its Applications* (J. L. Snell, ed.), CRC Press, Boca Raton.

Pemantle, R., Peres, Y.
1994 Planar first-passage percolation times are not tight, *Probability and Phase Transition* (G. R. Grimmett, ed.), Kluwer, Dordrecht, pp. 261–264.
1996a On which graphs are all random walks in random environments transient?, *Random Discrete Structures* (D. Aldous and R. Pemantle, eds.), IMA Volumes in *Mathematics and its Applications*, vol. 76, Springer, New York, pp. 207–211.
1996b Galton–Watson trees with the same mean have the same polar sets, *Annals of Probability* 23, 1102–1124.

Penrose, M.
1996 Continuum percolation and Euclidean minimal spanning trees in high dimensions, *Annals of Applied Probability* 6, 528–544.

Penrose, M. and Pisztora, A.
1996 Large deviations for discrete and continuous percolation, *Advances in Applied Probability* 28, 29–52.

Peres, Y.
1996 Intersection-equivalence of Brownian paths and certain branching processes, *Communications in Mathematical Physics* 177, 417–434.

Peyrière, J.
1978 Mandelbrot random beadsets and birthprocesses with interaction, I.B.M. Research Report RC-7417.

Pisztora, A.
1996 Surface order large deviations for Ising, Potts and percolation models, *Probability Theory and Related Fields* 104, 427–466.

Pokorny, M., Newman, C. M., Meiron, D.
1990 The trapping transition in dynamic (invasion) and static percolation, *Journal of Physics A: Mathematical and General* 23, 1431–1438.

Potts, R. B.
1952 Some generalized order-disorder transformations, *Proceedings of the Cambridge Philosophical Society* 48, 106–109.

Quas, A.
1996 Some properties of Lorentz lattice gas models, *Probability Theory and Related Fields* (to appear).

Reed, M., Simon, B.
1975 *Methods of Modern Mathematical Physics. II. Fourier Analysis, Self-Adjointness*, Academic Press, New York.

Reimer, D.
1997 Proof of the van den Berg–Kesten inequality (to appear).

Reynolds, P. J., Stanley, H. E., Klein, W.
1980 Large-cell Monte Carlo renormalization group for percolation, *Physical Review B* 21, 1223–1245.

Richardson, D.
1973 Random growth in a tesselation, *Proceedings of the Cambridge Philosophical Society* 74, 515–528.

Rogers, C. A.
1964 *Packing and Covering*, Cambridge University Press, Cambridge.

Roy, R.
1990 The Russo–Seymour–Welsh theorem and the equality of critical densities and the "dual" critical densities for continuum percolation on $\mathbb{R}^2$, *Annals of Probability* 18, 1563–1575.

Roy, R., Sarkar, A., White, D.
1998 Backbends in directed percolation, *Journal of Statistical Physics* 91, 889–908.

Rudin, W.
1962 *Fourier Analysis on Groups*, Wiley–Interscience, New York.

Ruelle, D.
1968 Statistical mechanics of a one-dimensional lattice gas, *Communications in Mathematical Physics* 9, 267–278.

Ruijgrok, T. W., Cohen, E. G. D.
1988 Deterministic lattice gas models, *Physics Letters A* 133, 415–418.

Rüschendorf, L.
1982 Comparison of percolation probabilities, *Journal of Applied Probability* 19, 864–868.

Russo, L.
1978 A note on percolation, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 43, 39–48.
1981 On the critical percolation probabilities, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 56, 229–237.
1982 An approximate zero–one law, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 61, 129–139.

Scheinerman, E. R., Wierman, J. C.
1987 Infinite AB percolation clusters exist, *Journal of Physics A: Mathematical and General* 20, 1305–1307.

Schonmann, R. H.
1998a Stability of infinite clusters in supercritical percolation, *Probability Theory and Related Fields* (to appear).
1998b Percolation in $\infty + 1$ dimensions at the uniqueness threshold (to appear).

Schulman, L. S.
1983 Long-range percolation in one dimension, *Journal of Physics A: Mathematical and General* 16, L639–L641.

Sevsek, F., Debierre, J. M., Turban, L.
1983 Antipercolation on Bethe and triangular lattices, *Journal of Physics A: Mathematical and General* 16, 801–810.

Seymour, P. D., Welsh, D. J. A.
1978 Percolation probabilities on the square lattice, *Advances in Graph Theory* (B. Bollobás, ed.), Annals of Discrete Mathematics 3, North-Holland, Amsterdam, pp. 227–245.

Shante, V. K. S., Kirkpatrick, S.
1971 An introduction to percolation theory, *Advances in Physics* 20, 325–357.

Simon, B.
1980 Correlation inequalities and the decay of correlations in ferromagnets, *Communications in Mathematical Physics* 77, 111–126.

Slade, G.
1991 The lace expansion and the upper critical dimension for percolation, *Mathematics of Random Media* (W. E. Kohler and B. S. White, eds.), Lectures in Applied Mathematics, vol. 27, AMS, Providence, R. I., pp. 53–63.
1995 Bounds on the self-avoiding walk connective constant, Special Issue: Proceedings of the Conference in Honor of Jean-Pierre Kahane, 1993, *Journal of Fourier Analysis and its Applications*, 525–533.

Smythe, R. T., Wierman, J. C.
1978 *First-Passage Percolation on the Square Lattice*, Lecture Notes in Mathematics no. 671, Springer, Berlin.

Spohn, H.
1991 *Large Scale Dynamics of Interacting Particles*, Springer, Berlin.

Stacey, A. M.
1996 The existence of an intermediate phase for the contact process on trees, *Annals of Probability* 24, 1711–1726.

Stanley, H. E.
1971 *Introduction to Phase Transitions and Critical Phenomena*, Clarendon Press, Oxford.

Stanley, H. E., Reynolds, P. J., Redner, S., Family, F.
1982 Position-space renormalization group for models of linear polymers, branched polymers, gels, *Real-Space Renormalization* (T. W. Burkhardt and J. M. J. van Leeuwen, eds.), Springer, Berlin, pp. 169–206.

Stauffer, D.
1979 Scaling theory of percolation clusters, *Physics Reports* 54, 1–74.
1981 Scaling properties of percolation clusters, *Disordered Systems and Localization* (C. Castellani, C. DiCastro, L. Peliti, eds.), Springer, Berlin, pp. 9–25.
1985 *Introduction to Percolation Theory*, Taylor and Francis, London.
1986 Percolation and cluster size distribution, *On Growth and Form* (H. E. Stanley and N. Ostrowsky, eds.), Nijhoff, Dordrecht, pp. 79–100.

Stauffer, D., Aharony, A.
1991 *Introduction to Percolation Theory*, 2nd Edition, Taylor and Francis, London.

Stinchcombe, R. B.
1974 Conductivity and spin-wave stiffness in disordered systems—an exactly soluble model, *Journal of Physics C: Solid State Physics* 7, 179–203.

Straley, J. P.
1977 Critical exponents for the conductivity of random resistor lattices, *Physical Review B* 15, 5733–5737.

Suzuki, M.
1965 A theorem on the exact solution of a spin system with a finite magnetic field, *Physics Letters* 18, 233–234.

Sykes, M. F., Essam, J. W.
1963 Exact critical percolation probabilities for bond and site problems in two dimensions, *Physical Review Letters* 10, 3–4.
1964 Exact critical percolation probabilities for site and bond problems in two dimensions, *Journal of Mathematical Physics* 5, 1117–1127.

Talagrand, M.
1995 Concentration of measure and isoperimetric inequalities in product spaces, *Institut des Hautes Etudes Scientifiques, Publications Mathématiques* 81, 73–205.

Tanemura, H.
1996 Critical behavior for a continuum percolation model, *Probability Theory and Mathematical Statistics (Tokyo 1995)*, World Scientific, pp. 485–495.

Tasaki, H.
1987a Geometric critical exponent inequalities for general random cluster models, *Journal of Statistical Physics* 49, 841–847.
1987b Hyperscaling inequalities for percolation, *Communications in Mathematical Physics* 113, 49–65.
1989 On the upper critical dimension of random spin systems, *Journal of Statistical Physics* 54, 163–170.

Tempel'man, A. A.
1972 Ergodic theorems for general dynamical systems, *Transactions of the Moscow Mathematical Society* 26, 94–132.

Temperley, H. N. V., Lieb, E. H.
1971 Relations between the 'percolation' and 'colouring' problem and other graph-theoretical problems associated with regular planar lattices; some exact results for the 'percolation' problem, *Proceedings of the Royal Society of London, Series A* 322, 251–280.

Tóth, B.

1985 A lower bound for the critical probability of the square lattice site percolation, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 69, 19–22.

Toulouse, G.

1974 Perspectives from the theory of phase transitions, *Il Nuovo Cimento B* 23, 234–240.

Vahidi-Asl, M. Q., Wierman, J. C.

1990 First-passage percolation on the Voronoi tesselation and Delaunay triangulation, *Random Graphs '87*, Wiley, London, pp. 341–359.

1990 A shape result for first-passage percolation on the Voronoi tesselation and Delaunay triangulation, *Random Graphs '89*, Wiley, London, pp. 247–262.

1993 Upper and lower bounds for the route length of first-passage percolation in Voronoi tesselations, *Bulletin of the Iranian Mathematical Society* 19, 15–28.

Varadhan, S. R. S.

1979 Boundary value problems with rapidly oscillating random coefficients, *Random Fields*, Colloquia of Mathematical Society János Bolyai, Esztergom, 27, North-Holland, Amsterdam, pp. 835–873.

Wang, F., Cohen, E. G. D.

1995 Diffusion in Lorentz lattice gas cellular automata: the honeycomb and quasi-lattices compared with the square and triangular lattices, *Journal of Statistical Physics* 81, 467–495.

Watts, G. M. T.

1996 A crossing probability for critical percolation in two dimensions, *Journal of Physics A: Mathematical and General* 29, L363–L368.

Welsh, D. J. A.

1986 Correlated percolation and repulsive particle systems, *Stochastic Spatial Processes* (P. Tautu, ed.), Lecture Notes in Mathematics no. 1212, Springer, Berlin, pp. 300–311.

1993 Percolation in the random-cluster process, *Journal of Physics A: Mathematical and General* 26, 2471–2483.

Wierman, J. C.

1978 On critical probabilities in percolation theory, *Journal of Mathematical Physics* 19, 1979–1982.

1981 Bond percolation on honeycomb and triangular lattices, *Advances in Applied Probability* 13, 293–313.

1982 Percolation theory, *Annals of Probability* 10, 509–524.

1984a A bond percolation critical probability determination based on the star–triangle transformation, *Journal of Physics A: Mathematical and General* 17, 1525–1530.

1984b Mixed percolation on the square lattice, *Journal of Applied Probability* 21, 247–259.

1985a Duality for directed site percolation, *Particle Systems, Random Media and Large Deviations* (R. Durrett, ed.), Contemporary Mathematics 41, AMS, Providence, R. I., pp. 363–380.

1985b Critical percolation probabilities, *Random Graphs '83* (M. Karoński, ed.), Annals of Discrete Mathematics 28, North-Holland, Amsterdam, pp. 349–359.

1987a Directed site percolation and dual filling models, *Annals of Discrete Mathematics* 33, 339–352.

1987b Duality for k-degree percolation on the square lattice, *Percolation Theory and Ergodic Theory of Infinite Particle Systems* (H. Kesten, ed.), IMA Volumes in *Mathematics and its Applications*, vol. 8, Springer, New York, pp. 311–323.
1988a Bond percolation critical probability bounds derived by edge contraction, *Journal of Physics A: Mathematical and General* 21, 1487–1492.
1988b AB percolation on close-packed graphs, *Journal of Physics A: Mathematical and General* 21, 1939–1944.
1988c On AB percolation in bipartite graphs, *Journal of Physics A: Mathematical and General* 21, 1945–1949.
1989 AB percolation: a brief survey, *Combinatorics and Graph Theory*, vol. 25, Banach Centre Publications, Warsaw, pp. 241–251.
1990 Bond percolation critical probability bounds for the Kagomé lattice by a substitution method, *Disorder in Physical Systems* (G. R. Grimmett and D. J. A. Welsh, eds.), Clarendon Press, Oxford, pp. 349–360.
1992 Equality of the bond percolation critical exponents for two pairs of dual lattices, *Combinatorics, Probability, Computing* 1, 95–105.
1994 Equality of directional critical exponents in multiparameter percolation models, *Journal of Physics A: Mathematical and General* 27, 1851–1858.
1995 Substitution method critical probability bounds for the square lattice site percolation model, *Combinatorics, Probability, Computing* 4, 181–188.

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

Wilkinson, D., Willemsen, J. F.
1983 Invasion percolation: a new form of percolation theory, *Journal of Physics A: Mathematical and General* 16, 3365–3376.

Willemsen, J. F.
1984 Investigations on scaling and hyperscaling for invasion percolation, *Physical Review Letters* 52, 2197–2200.

Wilson, R. J.
1979 *Introduction to Graph Theory*, Longman, London.

Wu, F. Y.
1982 The Potts model, *Reviews in Modern Physics* 54, 235–268.

Yang, W., Zhang, Y.
1992 A note on differentiability of the cluster density for independent percolation in high dimensions, *Journal of Statistical Physics* 66, 1123–1138.

Zhang, Y.
1988 Unpublished.
1991 A power law for connectedness of some random graphs at the critical point, *Random Structures and Algorithms* 2, 101–119.
1992 Failure of the power laws on some subgraphs of the Z^2 lattice, *Journal of Physics A: Mathematical and General* 25, 6617–6622.
1993 A shape theorem for epidemics and forest fires with finite range interactions, *Annals of Probability* 21, 1755–1781.
1994a A note on inhomogeneous percolation, *Annals of Probability* 22, 803–820.
1994b Analyticity properties at the supercritical state, preprint.
1995a The fractal volume of the two-dimensional invasion percolation cluster, *Communications in Mathematical Physics* 167, 237–254.

1995b Supercritical behaviors in first-passage percolation, *Stochastic Processes and their Applications* 59, 251–266.
1995c A limit theorem for matching random sequences allowing deletions, *Annals of Applied Probability* 5, 1236–1240.
1996a The complete convergence theorem on trees, *Annals of Probability* 24, 1408–1443.
1996b Continuity of percolation probability in $\infty + 1$ dimensions, *Journal of Applied Probability* 33, 427–433.
1996c Divergence of the bulk resistance at criticality in disordered media, *Journal of Statistical Physics* 84, 263–267.
1997a Two critical behaviors of first passage time, preprint.
1997b Some power laws on two dimensional critical bond percolation, preprint.
1998 A martingale approach in the study of percolation clusters on the Z^d lattice, preprint.

Zhang, Y., Zhang, Y. C.
1984 A limit theorem for N_{0n}/n in first passage percolation, *Annals of Probability* 12, 1068–1076.

Ziff, R. M., Kong, X. P., Cohen, E. G. D.
1991 Lorentz lattice-gas and kinetic-walk model, *Physical Review A* 44, 2410–2428.

Zuev, S. A.
1987 Percolation threshold bounds for the square lattice, *Theory of Probability and its Applications* (in Russian) 32, 606–609 (551–553 in translation).

Zuev, S. A., Sidorenko, A. F.
1985a Continuous models of percolation theory I, *Theoretical and Mathematical Physics* (in Russian) 62, 76–86 (51–58 in translation).
1985b Continuous models of percolation II, *Theoretical and Mathematical Physics* (in Russian) 62, 253–262 (171–177 in translation).

Index of Names

Subject Index