PERCOLATION

Geoffrey R. Grimmett

1. A brief history of percolation theory

A child of the 1950s, percolation theory has grown to mature adulthood over the intervening 45 years. It lies at the heart of an intense development within probability theory directed at a coherent theory of 'random spatial processes'. It finds applications in all areas of science, while continuing to provide a source of beautiful and provoking problems for mathematicians and physicists.

Following the presentation by Hammersley and Morton of a paper [40] on Monte Carlo methods to the Royal Statistical Society in 1954, Simon Broadbent contributed the following discussion [17]:

> "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.
>
> It is not difficult to express this problem for a finite lattice in a form suitable for Monte Carlo work by an electronic computer. The capacity of computers is, however, insufficient for any but very small lattices. That is another example of the authors' remark that pen and paper might be better than machine work. . . ."

These words were stimulated by Broadbent's work at the British Coal Utilization Research Association, where he was involved in the design of gas masks for coal miners. One of the authors of the RSS paper was John Hammersley[1], and he

1991 *Mathematics Subject Classification.* 60K35, 82B43.
This version was prepared on 15 September 1997.

[1]Hammersley is also the author of [39], from which some of this historical material is drawn.

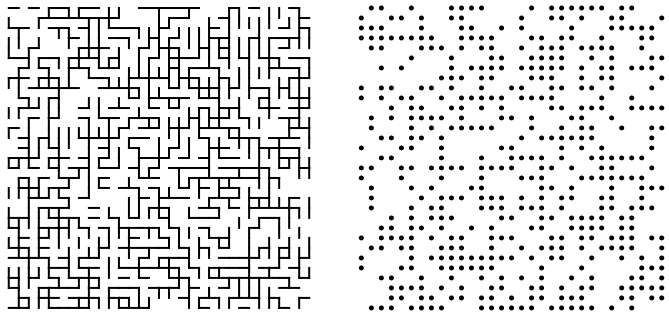


Fig. 1.1. Illustrations of bond and site percolation on the two-dimensional square lattice, with $p = \frac{1}{2}$.

recognised the potential of Broadbent's model[2]. Their subsequent collaboration led to a clear formulation of the percolation model, and to a striking series of early papers containing several of the principal methods to be discovered.

Here is the model in its basic form. We start with a 'crystalline lattice', and for the sake of simplicity we shall consider here only the hypercubic lattice $\mathbb{L}^d = (\mathbb{Z}^d, \mathbb{E}^d)$ in d dimensions. (The vertex set $\mathbb{Z}^d$ is the set of all d-vectors of integers, and the edge set $\mathbb{E}^d$ contains all unordered pairs $\langle x, y\rangle$ of vertices $x, y \in \mathbb{Z}^d$ separated by unit Euclidean distance.) Let $0 \le p \le 1$, and suppose we are provided with a coin which shows heads on each toss with probability p. We flip this coin once for each edge e, and we call e *open* if the coin shows heads; the edge e is called *closed* if tails are shown. The outcome of the entire experiment is a subgraph of $\mathbb{L}^d$ having vertex set $\mathbb{Z}^d$ together with all open edges. *Bond percolation theory* is the theory of the geometry of this open graph. See Figure 1.1 for an illustration of two-dimensional bond percolation with $p = \frac{1}{2}$.

The words 'open/closed' indicate that each edge is in one of exactly two available states, and these words have an appealing physical motivation. We may think of an open edge as being open to the transmission of fluid, and of a closed edge as being blocked. If fluid is supplied at a given vertex x, then it wets exactly the set C_x of vertices y having the property that there exists a path of open edges from x to y. We call C_x the *open cluster at* x, and we write $C = C_0$ for the open cluster at the origin 0.

There are many physical situations to which the percolation model is relevant, of which the following is a naive example. A porous stone is immersed in a bucket of water. What is the probability that water reaches the centre of the stone? We may choose to model a porous stone as a large finite subset S of the lattice $\mathbb{L}^d$. Assuming that water flows along open edges but not along closed edges, we are asked to calculate the probability that the centre of the stone (the origin of $\mathbb{L}^d$, say) is joined by a path of open edges to some vertex on its surface ∂S, or alternatively that $C \cap \partial S \neq \varnothing$. If S is large, or, alternatively expressed, the structure of edges is

[2]Hammersley and Morton replied with foresight to Broadbent's discussion: "Mr. Broadbent's problem is very fascinating and difficult . . .".

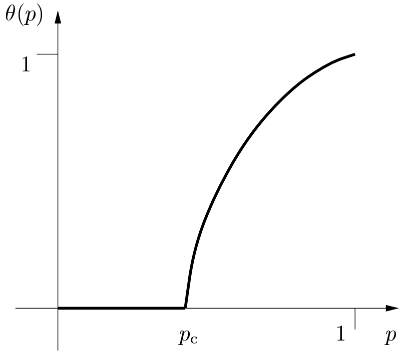


Fig. 1.2. A sketch of the percolation probability $\theta(p)$. Not all the obvious features of this function have been proved rigorously. Note the existence of a critical point p_c. Perhaps the principal open conjecture is that θ is continuous at the critical point p_c in all dimensions; this has been proved so far only when either $d = 2$ or $d \geq 19$.

exceedingly fine, then this probability is close to the probability that C is infinite. This crude physical argument leads to the central question of percolation theory: what is the probability $\theta(p)$ that the origin lies in an *infinite* open cluster? Subject to an appropriate interpretation of Broadbent's model, the quantity $\theta(p)$ is exactly the function $\Lambda(p)$ occurring in the quotation at the beginning of this section. The quantity $\theta(p)$ is called the *percolation probability*, and is sketched in Figure 1.2.

An exact calculation of $\theta(p)$ seems inconceivable. The marriage of *geometry* and *probability* is challenging and often uncomfortable; exact results are rare, and exist apparently only when there exists special structure. However, many properties of the function θ have been discovered. Its principal property is that of *phase transition*. Clearly θ is non-decreasing, since the addition of open edges may create infinite paths but cannot destroy them. Therefore there exists a *critical value* p_c for p, defined by the statement that

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

The fundamental property of percolation is that p_c is non-trivial if $d \geq 2$, which is to say that

$$0 < p_c < 1 \qquad \text{if } d \geq 2.$$

This was proved in [18, 37, 38]; the proof is so important 'beyond percolation' that some details will be presented in Section 2.1. (It is elementary that $p_c = 1$ if $d = 1$, and we assume henceforth that $d \geq 2$.)

One of the principal targets of modern probability theory and statistical physics is *to understand phase transitions and critical phenomena*. Although the physical theory is largely well developed and widely accepted, the rigorous mathematical theory contains major open challenges. The percolation model contains a maximum of (statistical) independence, and has proved a superb testing ground for new

methodology. Notwithstanding the large amount of effort expended on the percolation phase transition, many of the central questions remain opaque (see Section 2.4). For example, it is unknown whether $\theta(p_c) = 0$ or $\theta(p_c) > 0$ in general, although it is widely believed that $\theta(p_c) = 0$. The corresponding property of branching processes, namely that a critical branching process is almost surely finite, is fully understood and relatively elementary. A proof that $\theta(p_c) = 0$ for all $d \geq 2$ would, on the other hand, answer a long-standing open question, and will probably require new ideas. The interested reader is challenged to prove that $\theta(p_c) = 0$ when $d = 3$. We note that $\theta(p_c) = 0$ was proved by Harris [44] and Kesten [49] when $d = 2$, and by Hara and Slade [42, 43] when d is sufficiently large ($d \geq 19$ is certainly enough).

Percolation theory has earned a reputation as a source of hard problems which are easy to state and whose solutions require new methods. The most provocative such problem was the conjecture that $p_c = \frac{1}{2}$ for bond percolation on the square lattice $\mathbb{L}^2$. Originally conjectured around 1955, the simplicity of the statement provoked many to attempt a solution. In a beautiful paper [44] dated 1960, Theodore Harris proved that $\theta(\frac{1}{2}) = 0$ for $\mathbb{L}^2$, thereby deriving that $p_c \geq \frac{1}{2}$. Numerical simulations suggested that p_c was a little less than $\frac{1}{2}$, and what better evidence could support the conjecture? When, in 1963, Sykes and Essam announced a solution to this and related problems, much interest was aroused ([70, 71]). Unfortunately their arguments, although reasonable, lacked a totally rigorous foundation. (Even today, we are unable to confirm or deny a key hypothesis of their approach.)

Percolation theory entered a period of recession for mathematicians, from which it emerged in 1978 with the simultaneous and independent publications of papers by Russo [67] and Seymour and Welsh [68] devoted to bond percolation on the square lattice. This was the spur to Harry Kesten which led to his beautiful proof ([49]) that $p_c = \frac{1}{2}$ for the square lattice. A masterpiece of probabilistic and geometrical argument, this theorem was the beginning of a percolation era of vigour and richness.

What is the rationale for this exact calculation in two dimensions? To a planar graph G we may associate a planar dual G_d constructed by placing a vertex inside each face of G, and by joining two such vertices by a dual edge e_d whenever the corresponding faces of G share a boundary edge e. Now consider a bond percolation process on an infinite planar graph G. This induces a percolation process on the dual graph G_d according to the rule: a dual edge e_d is *open* if and only if the corresponding primal edge e is *closed*. It may be seen (as in Figure 1.3) that $|C_x| < \infty$ if and only if the vertex x lies within the interior of some circuit of open edges of the dual graph G_d. This observation, taken together with the fact that the dual of the square lattice is isomorphic to the square lattice, is at the heart of Harris's result that $\theta(\frac{1}{2}) = 0$ for the square lattice. Kesten's proof of $p_c = \frac{1}{2}$ exploited further the self-duality of $\mathbb{L}^2$.

For a two-dimensional lattice $\mathcal{L}$ and its dual $\mathcal{L}_d$, one may sometimes show that

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

whence $p_c(\mathcal{L}) = \frac{1}{2}$ for appropriate self-dual lattices $\mathcal{L}$. In the presence of another link between $\mathcal{L}$ and $\mathcal{L}_d$, then an exact calculation may sometimes be derivable. For example, the dual of the triangular lattice is the hexagonal lattice (see Figure 1.4).

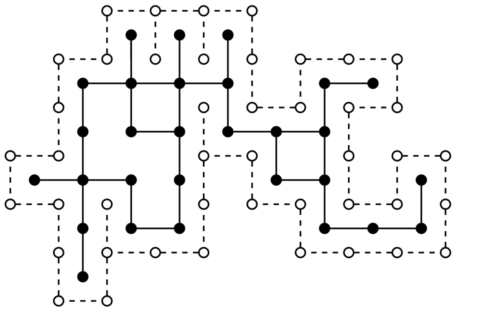


Fig. 1.3. A finite cluster of the square lattice, surrounded by an open dual circuit of the dual lattice.

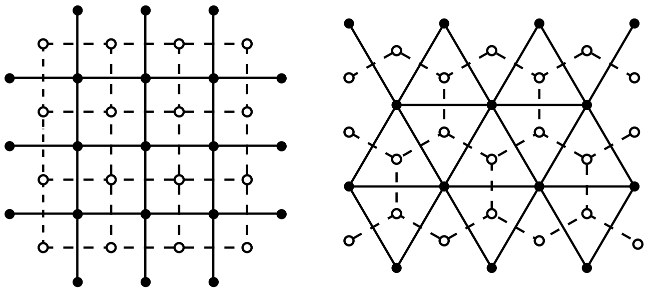


Fig. 1.4. The square lattice is self-dual. The dual of the triangular lattice is the hexagonal lattice.

The so-called 'star–triangle transformation' provides another link between these two lattices, and one may then conclude that the triangular lattice has critical probability $2\sin(\pi/18)$ and the hexagonal lattice $1-2\sin(\pi/18)$. See [71, 72].

However beautiful these exact calculations, they mark exceptions rather than rules. In the absence of an argument such as duality, there seems no reason to expect percolation quantities to be calculable. A vast amount of effort and ingenuity has been invested in deriving numerical estimates of such quantities, especially of critical probabilities. There is a variety of methods in use, from pure 'Monte Carlo' to the partly analytical, and the modern computer has enabled a reasonable degree of accuracy; see [47, p. 175] for example. However, in most cases, the corresponding rigorous upper and lower bounds are quite far apart.

In bond percolation, the randomness is associated with the *edges* of the lattice. If, instead, each *vertex* is designated open or closed at random, then the ensuing model is termed *site percolation* (illustrated in Figure 1.1). One may consider also 'mixed' models in which both edges and vertices are given random states. Indeed there is a multiplicity of possible generalisations.

In another variant, called *oriented* (or *directed*) *percolation*, some or all of the

edges of the lattice are assigned a particular orientation, and one asks whether or not there exists an infinite open path from the origin which conforms to the orientations.

We represent the probability function by P_p, and expectation by E_p. The letters 'a.s.' are an abbreviation for 'almost surely', and mean that the corresponding statement has probability 1. The Euclidean norm on $\mathbb{R}^d$ is denoted as $|\cdot|$.

The next section contains an account of the current mathematical theory of percolation, placed in a historical perspective. Ideas from percolation have proved to be of major importance in studying a variety of disordered systems, and Sections 3–6 contain thumbnail sketches of just a few of these, namely first-passage percolation, epidemic models, a 'Lorentz lattice gas', and ferromagnetism via the random-cluster model.

The principal mathematical accounts of percolation are [28, 31, 50], and other books include [1, 26, 47]. No serious attempt has been made here to provide a comprehensive list of references.

2. The mathematical theory

We explore next certain themes of the rigorous theory, and attempt to summarise the principal progress as well as future directions for research.

2.1 Existence of phase transition

It is a fundamental fact that the critical probability p_c satisfies the strict inequalities $0 < p_c < 1$, so long as the number d of dimensions satisfies $d \geq 2$. The proof is based on simple but beautiful ideas, and is canonical in the sense of providing a template for proving the existence of critical phenomena in a range of disordered systems.

In proving that $p_c > 0$, one uses the idea of a self-avoiding walk (SAW). A SAW is a path of the lattice which visits no vertex more than once. Let f_n be the number of SAWs on the lattice $\mathbb{L}^d$ having n edges and with the origin as an endpoint. In their famous paper [40] referred to above, Hammersley and Morton showed the subadditive relation $f_{m+n} \leq f_m f_n$, whence the exponential asymptotic $f_n = \kappa^{n+o(n)}$ follows for some constant $\kappa = \kappa(d)$ called the *connective constant* of the lattice. It may be shown that $1 < \kappa < 2d-1$. (By the way, it is a problem of great appeal to understand the behaviour of the error term $o(n)$. Indeed the theory of self-avoiding walks is a first cousin of percolation theory, and this last problem has an exact analogue in percolation. For accounts of the modern theory of SAWs, see [60].)

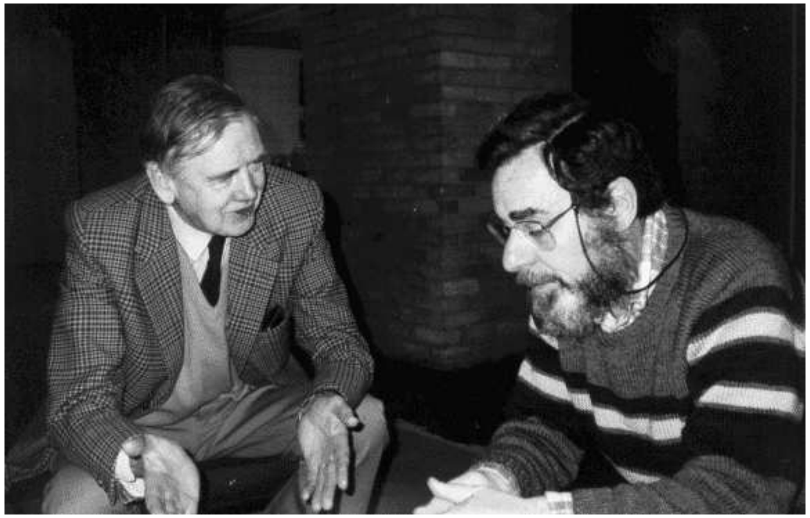
Let N_n be the number of open self-avoiding walks (i.e., SAWs of open edges) with length n and having the origin as an endpoint. If $|C| = \infty$, then $N_n \geq 1$ for all n, whence

$$\theta(p) \leq P_p(N_n \geq 1) \leq E_p(N_n) = f_n p^n. \tag{2.1}$$

Since $f_n = \kappa^{n+o(n)}$, we deduce that $\theta(p) = 0$ if $p\kappa < 1$, whence $p_c \geq \kappa^{-1}$ as required.

In proving that $p_c < 1$, we note first that $p_c = p_c(\mathbb{L}^d)$ is non-increasing in d (since $\mathbb{L}^d$ may be viewed as a subgraph of $\mathbb{L}^{d+1}$). Therefore it will suffice to prove

Harry Kesten, Rudolf Peierls, and Roland Dobrushin in the Front Quadrangle of New College, Oxford, November 1993.



John Hammersley and Harry Kesten in the Mathematical Institute, Oxford, November 1993.

$p_c < 1$ for the square lattice $\mathbb{L}^2$. To achieve this we require an additional idea, namely that of two-dimensional duality. Such a method as the following is often called a 'Peierls argument', after Rudolf Peierls who made use of it in his proof of the existence of a phase transition for the Ising model of ferromagnetism ([63]).

Let G_n be the number of open circuits in the dual lattice of $\mathbb{L}^2$ having length n and with the origin of $\mathbb{L}^2$ in their interior. As noted above, C is finite if and only if $G_n \geq 1$ for some n. Therefore,

$$P_p(|C| < \infty) \leq \sum_n P_p(G_n \geq 1) \leq \sum_n E_p(G_n) = \sum_n g_n(1-p)^n,$$

where g_n is the number of 'self-avoiding' circuits of the dual lattice having length n and with the origin in their interior. We require an upper bound for g_n. It is 'easy' to see that $g_n \leq n f_{n-1}$, and that $g_n = 0$ if $n \leq 3$. Hence,

$$1 - \theta(p) \leq \sum_{n=4}^{\infty} n\kappa^{n+o(n)}(1-p)^n.$$

The summation may be made strictly smaller than 1 by a sufficiently small (but positive) choice of $1-p$. For such p, we have that $\theta(p) > 0$, whence $p_c < 1$. With a little extra work, one may obtain that $(1-p_c)\kappa \geq 1$, which is to say that $p_c \leq 1 - \kappa^{-1}$.

Since $0 < p_c < 1$, a percolation model has a *subcritical* phase (when $p < p_c$), a *supercritical* phase (when $p > p_c$), and a *critical point* (when $p = p_c$). The subcritical and supercritical phases are now fairly well understood. In contrast, there are substantial open questions concerning the nature of the phase transition. The physical picture provided by so called 'scaling theory' is widely accepted by both mathematicians and physicists, but it remains a major challenge to mathematicians to provide a rigorous foundation. The following three subsections are devoted respectively to the subcritical and supercritical phases, and to the critical behaviour of percolation.

2.2 The subcritical phase

We define the *radius* of the open cluster C at the origin by $\mathrm{rad}(C) = \max\{|x| : x \in C\}$. Already in [37], Hammersley sought a proof of 'exponential decay' when $p < p_c$, or more precisely that

$$P_p\big(\mathrm{rad}(C) \geq n\big) \leq e^{-n\alpha(p)} \qquad \text{for } n \geq 1, \text{ where } \alpha(p) > 0. \tag{2.2}$$

Exponential decay turns out to be the key to understanding the subcritical phase. It provides a tool for estimating the probabilities of large open clusters with given properties. For example, it implies that the largest open cluster intersecting the cube $[-n,n]^d$ has cardinality of order $\{d/\alpha(p)\}\log(2n)$, for large n.

Exponential decay turns out to be linked to the superficially weaker statement that the mean cluster size

$$\chi(p) = E_p|C| = \sum_{n=1}^{\infty} nP_p(|C| = n)$$

satisfies

$$\chi(p) < \infty \qquad \text{when } p < p_c. \tag{2.3}$$

It is not difficult to show that exponential decay implies $\chi(p) < \infty$, and Hammersley [37] demonstrated the converse as early as 1957.

There has been a concrete objective for proving that $\chi(p) < \infty$ when $p < p_c$. Using delicate path-intersection properties of $\mathbb{L}^2$, Russo, Seymour, and Welsh [67, 68] proved in 1978 that $p_c(\mathbb{L}^2) = \frac{1}{2}$ if (2.3) holds for $\mathbb{L}^2$. This provoked Kesten's proof [49] of (2.3) for the square lattice, and hence his famous exact calculation of p_c.

The natural generalisation of Kesten's theorem was therefore (2.3) for general lattices, or equivalently that exponential decay is valid for $p < p_c$. Two separate proofs of this emerged in 1986, from independent and distinct sources. At a conference in Minneapolis, Aizenman and Barsky [4] announced a family of differential inequalities whose solution yielded (2.3). A special feature of their method was that it was capable of adaptation to other models; indeed the Ising model has been studied similarly [2, 5].

Almost simultaneously in Moscow, and working in relative isolation, Menshikov proved exponential decay when $p < p_c$, by a direct argument. In those days, communication between East and West was fragmented and problematic. Aided by the First World Congress of the Bernoulli Society, held in Tashkent in September 1986, Menshikov's results reached the West, albeit in Russian.

The 1980s were a wonderful period for percolation. The decade opened with the long awaited resolution of the conjecture that $p_c(\mathbb{L}^2) = \frac{1}{2}$. Through the subsequent work of Aizenman, Barsky, Chayes, Chayes, Kesten, Menshikov, Newman and their colleagues, the general theory developed to encompass more general questions and dimensions, and then exponential decay emerged in total generality. At the end of the decade, the principal structure of the supercritical phase was established, as summarised next.

2.3 The supercritical phase

Suppose that p is such that $\theta(p) > 0$, and let I denote the number of infinite open clusters. There is an easy argument using Kolmogorov's zero–one law which implies that $P_p(I \geq 1) = 1$, which is to say that there exists a.s. an infinite open cluster. Could there exist two or more such clusters? Intuition suggests not, within the confines of a finite-dimensional space, but proof was required. Further stimulation for the question was provided by the realisation ([14]) that θ is continuous at any point p ($> p_c$) so long as $I = 1$ a.s. for that value of p.

The a.s. uniqueness of the infinite open cluster was finally established in 1987 ([7]), by an arguably mysterious method involving 'boundary conditions' and a quantitative estimate of 'large deviation' type. Any mystery was removed by the subsequent work of Burton and Keane [21], who decoupled the geometry from the probability in a transparent manner, thus achieving a proof of uniqueness which uses no estimate, but only the 'ergodicity' of the 'product measure' P_p. Their beautiful argument proceeds by the following sequence of steps.

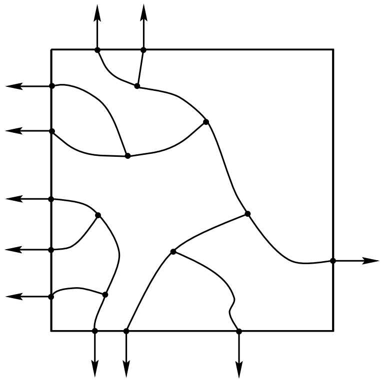


Fig. 2.1. An illustration of the map of trifurcations inside a box. Note the forest-like structure.

Let us call a vertex x a *trifurcation* if: $|C_x| = \infty$, but the removal of x turns C_x into three disjoint infinite open clusters and no finite clusters. Suppose henceforth that $0 < p < 1$.

1. *There exists a constant m, depending on p, such that $P_p(I = m) = 1$.* (Proof by ergodicity.)
2. *It must be the case that $m \in \{0, 1, \infty\}$.* (Proof by contradiction: Suppose that $m \geq 2$. Find a large box B which intersects two or more infinite open clusters. By making all edges within B open, we may obtain that $P_p(I \leq m - 1) > 0$. This is a contradiction unless $m = m - 1$, which is to say that $m = \infty$.)
3. Suppose $m = \infty$ (so that $m \geq 3$, in particular). *It follows that $\pi = P_p(0 \text{ is a trifurcation}) > 0$.* (Proof: Find a large box B which intersects three or more infinite open clusters, and 're-define' the states of edges inside B in such a way that the origin becomes a trifurcation.)
4. *It is the case that $m \in \{0, 1\}$.* (There follows the geometrical part of the proof. Suppose that $m = \infty$. By Step 3 and the ergodic theorem, the number of trifurcations inside the box $B_n = [-n, n)^d$ has order of magnitude $\pi|B_n|$. We draw a map of these trifurcations, and the paths between them (see Figure 2.1), thereby obtaining a forest-like graph of degree 3 having boundary vertices lying in the surface ∂B_n of B_n. It is an elementary fact of graph theory that such graphs have boundary comparable in size to their volume, whence $|\partial B_n| \geq c\pi|B_n|$ for some $c > 0$. However $|\partial B_n|$ and $|B_n|$ have orders of magnitude $(2n)^{d-1}$ and $(2n)^d$ respectively. This provides a contradiction for large n.)

The above argument is well suited for generalisations to other models, and it has proved extremely useful in other contexts (see [16, 30] for example).

If the finiteness of $\chi(p)$ is the key to the subcritical phase, what is the corresponding key to the supercritical phase? Using duality, one may see that two-dimensional bond percolation is supercritical if and only if its dual is subcritical. This fact enables a fairly full study of the supercritical phase in two dimensions. However, the picture is much more complicated when $d \geq 3$.

Suppose $d \geq 3$. Let M be a positive integer, and let $S_M = \{0, 1, \dots, M\}^{d-2} \times \mathbb{Z}^2$ be a 'slab' of $\mathbb{L}^d$ having thickness M. Interpreting S_M as a subgraph of $\mathbb{L}^d$, then S_M has a critical probability $p_c(S_M)$. Since $S_M \subseteq S_{M+1} \subseteq \mathbb{Z}^d$, we have that $p_c(S_M) \geq p_c(S_{M+1}) \geq p_c$; therefore the limit $\widetilde{p}_c = \lim_{m\to\infty} p_c(S_M)$ exists and satisfies $\widetilde{p}_c \geq p_c$. We now ask whether or not

$$\widetilde{p}_c = p_c. \tag{2.4}$$

Given that $\widetilde{p}_c = p_c$, one may study in detail the supercritical phase. More precisely, if p satisfies $p > \widetilde{p}_c$, then much may be learned about the corresponding percolation model, by exploiting and refining the following rough argument. If $p > \widetilde{p}_c$ then $p > p_c(S_M)$ for some M. Now, $\mathbb{Z}^d$ may be partitioned into translates of S_M. Each such translate is (topologically) two-dimensional, and is supercritical (since $p > p_c(S_M)$). Two such translates are disjoint, and therefore the corresponding percolation processes are independent. It follows that, a.s., each translate of S_M possesses an (essentially) two-dimensional infinite open cluster. If required, one may obtain estimates for the geometry of such a cluster by using two-dimensional arguments. In this way, one gains a control over the geometry of percolation in $\mathbb{Z}^d$, whence estimates of value follow. If $\widetilde{p}_c = p_c$, then such estimates are valid throughout the supercritical phase.

Here are some examples of possible conclusions. Unlike the subcritical phase, the decay of $P_p(|C| = n)$ is not exponential. Rather, there exist positive functions $\alpha(p), \beta(p)$ such that

$$\exp\big(-\alpha(p)n^{(d-1)/d}\big) \leq P_p(|C| = n) \leq \exp\big(-\beta(p)n^{(d-1)/d}\big),$$

which is to say that the decay is 'stretched exponential'. It is believed that the limit

$$\lim_{n\to\infty} \left\{ -\frac{1}{n^{(d-1)/d}} \log P_p(|C| = n) \right\}$$

exists, but this is known only when $d = 2$ (see [10]).

The above estimate concerns the *volume* of C. Turning to a one-dimensional measure of C, we define its *radius* $\mathrm{rad}(C) = \max\{|x| : x \in C\}$ as before. Using slab arguments as above ([23, 28, 31]), one may obtain that

$$P_p\big(n \leq \mathrm{rad}(C) < \infty\big) = e^{-n\gamma(p)+\mathrm{o}(n)}$$

for some $\gamma(p)$ satisfying $0 < \gamma(p) < \infty$.

Somewhat similar to (2.3) in the subcritical case, a completely separate argument is necessary for the proof that $\widetilde{p}_c = p_c$, and this was provided in 1990 by Grimmett and Marstrand [33]. Their proof uses what is known as a 'block' or 'renormalisation' argument, and such methods have proved very valuable in a variety of contexts. The basic idea is as follows. We partition $\mathbb{Z}^d$ into a union of (disjoint) translates of the cube $[0, M]^d$. We now construct a certain event for any given block with the

property that this event has probability close to 1 (for example, this event might be of the type 'there exist at least R open crossings of the block, in each of the d possible directions'). Viewing each block as a composite vertex in a new lattice, one obtains that the good blocks (i.e., the blocks for which the event in question occurs) form a supercritical site percolation process having density close to 1. One now combines known facts about such a 'renormalised' process, together with an appropriate choice of the event in question, in order to obtain properties of the original system. Various difficulties arise in developing this programme, but these may largely be overcome, thereby obtaining amongst other things that $\widetilde{p}_c = p_c$.

The absence of a general proof that $\theta(p_c) = 0$ was remarked earlier. Some impact on this question has been made by block arguments, but they are curiously incomplete (see [11, 33]). The following 'absurd' possibility has not been ruled out for $2 < d < 19$: when $p = p_c$, there exists a.s. an infinite open cluster in $\mathbb{Z}^d$, but no half-space of $\mathbb{Z}^d$ contains such a cluster.

2.4 At and near the critical point

It is one of the most fascinating problems of modern probability theory to build a rigorous theory of phase transitions. Percolation has been at the forefront of progress in recent years, but the story is far from complete. The basic question is to understand the behaviour of the process within a finite box, in the double limit as $p \to p_c$ and as the size of the box tends to infinity. A rich picture has emerged from the physics literature; mathematicians' understanding of this picture is significant but far from exhaustive (see [28, Chaps 7–8]).

Here is a mathematician's sketch of the physical theory. The critical point p_c marks a *singularity*, and otherwise smooth functions behave singularly at this point. It is believed that this singularity is of 'power' type. More precisely, it is believed that there exist non-trivial *critical exponents* β, γ such that

$$\begin{aligned} \theta(p) &\approx (p - p_c)^{\beta} && \text{as } p \downarrow p_c, \\ \chi(p) &\approx |p - p_c|^{-\gamma} && \text{as } p \uparrow p_c. \end{aligned}$$

(The asymptotic relation $\approx$ should be interpreted in an appropriate manner, for example in the manner of 'logarithmic asymptotics'.) In fact, any 'macroscopic quantity' should have a power law singularity; one may postulate thus at least five critical exponents, of which β and γ are but two.

Now fix $p = p_c$, and look on increasing 'length scales'. It is believed that there exist further critical exponents δ, ρ such that

$$\begin{aligned} P_{p_c}(|C| = n) &\approx n^{-1-1/\delta} && \text{as } n \to \infty, \\ P_{p_c}\big(\mathrm{rad}(C) = n\big) &\approx n^{-1-1/\rho} && \text{as } n \to \infty. \end{aligned}$$

Another exponent η is postulated in a similar manner.

Appealing but non-rigorous methods suggest that these eight critical exponents satisfy a collection of four 'scaling relations'. Further arguments suggest that, if d is not too large, then they satisfy two further relations called 'hyperscaling relations'.

It is generally the case that, for a given model of statistical mechanics, eight numbers can be defined in manners analogous to those alluded to above for percolation. Although these numbers will generally depend on the model, *they are expected to satisfy the same scaling and hyperscaling relations.* In addition, for any given model and number d of dimensions, their values should not depend on the particular lattice in use. For example, the values of β and δ should be the same for both bond and site percolation on both the square and triangular lattices. Such statements are referred to under the title of 'universality'.

Furthermore, for a two-dimensional percolation model, some individuals find reason to believe that $\beta = \frac{5}{36}$, $\gamma = \frac{43}{18}$, $\delta = \frac{91}{5}$, and so on. Such predictions are so distant from mathematical rigour that mathematicians tend to be shy of words such as 'believe' and 'accept' in this context.

Just as $d = 2$ is special, so is the case of 'large d'. The idea is that, when d is large, then the singularity should have the same qualities as when the lattice is replaced by a binary tree. Percolation on a tree is an old friend, namely a branching process. Exact calculations for a branching process lead to the prediction of exact values for critical exponents 'for large d', namely the 'mean-field values' $\beta = 1$, $\gamma = 1$, $\delta = 2$, and so on. In a remarkable series of papers, Hara and Slade [42, 43] have proved results of this type, based in part on work of Aizenman and Newman [8] and others. Their method is known as the 'lace expansion', and they have wielded it with virtuosity in their solutions to many models for large d, including lattice animals, self-avoiding walks, and percolation.

In order to be concrete, we state here some of the results of Hara and Slade: they have proved that β and γ exist and satisfy $\beta = \gamma = 1$, under the assumption that $d \geq 19$. It is believed that such calculations may be extendable to values of d satisfying $d \geq 7$, or even perhaps $d \geq 6$. For $2 \leq d < 6$, critical exponents are not expected to take their mean-field values.

One of the most remarkable families of conjectures of stochastic geometry to have emerged recently is that of 'conformal invariance'. It concerns two-dimensional percolation, and is supported by numerical evidence (see [57]). Roughly speaking, part of the conjecture is that crossing probabilities for *critical percolation* are invariant under conformal maps of $\mathbb{R}^2$. More precisely, let C be a simple closed curve of $\mathbb{R}^2$, and let α and β be arcs of C. Let $r > 0$, and consider the probability $P_{p_c}(r\alpha \leftrightarrow r\beta \text{ in } rC)$ that the interior of the dilated copy rC of C contains an open path joining the dilated arcs $r\alpha$ and $r\beta$. First, it is believed that the limit

$$\pi(\alpha, \beta; C) = \lim_{r\to\infty} P_{p_c}(r\alpha \leftrightarrow r\beta \text{ in } rC) \tag{2.5}$$

exists for all α, β, C.

Now, let ϕ be a conformal mapping on the interior C which is bijective up to its boundary. The principle of conformal invariance predicts that

$$\pi(\phi\alpha, \phi\beta; \phi C) = \pi(\alpha, \beta; C).$$

Extensive numerical simulation supports this conjecture. For fuller discussion of this principle and for recent results, see [3, 57].

A prospective relationship with conformal field theory had led to a startling prediction for exact values of crossing probabilities known as Cardy's formula [22].

For simplicity, let us consider site percolation on the square lattice $\mathbb{L}^2$. Take C to be a rectangle with side-lengths a and b, and consider the event that there exists an open crossing between its opposite sides having length a. When $p = p_c$, the corresponding quantity $\pi(a, b; C)$ given in (2.5) is conjectured to equal

$$\frac{3\Gamma(\frac{2}{3})}{\Gamma(\frac{1}{3})^2} \sin^{2/3}\theta \; {}_2F_1(\tfrac{1}{3}, \tfrac{2}{3}, \tfrac{4}{3}, \sin^2\theta)$$

where Γ is the gamma function, ${}_2F_1$ is a hypergeometric function, and θ is a known function of the ratio a/b.

3. First-passage percolation

First-passage percolation is the half-brother of percolation. It was formulated by Hammersley and Welsh [41] in 1965 as a time-dependent model for the flow of liquid through a porous body. Motivated by a need to understand the concept of the *velocity* of this flow, Hammersley and Welsh were led to the idea of a 'subadditive stochastic process'. Subadditive processes are now a standard tool of much power in probability theory.

We begin with the lattice $\mathbb{L}^d$ where $d \geq 2$. To each edge e we assign a random variable $T(e)$ (called the *time coordinate* of e), which we interpret as the time required for liquid to traverse e; we assume that the $T(e)$ are non-negative and independent, with some common distribution function F.

For any path π, the corresponding *passage time* $T(\pi)$ is the sum of the time coordinates of the edges in π. The *first-passage time* $a(x, y)$ between two vertices x and y is defined as the infimum of the passage times of all paths from x to y. If we supply liquid at x, then it will arrive at y after an elapsed time $a(x, y)$.

How fast does liquid spread through the medium? It is a basic observation that the $a(x, y)$ satisfy the subadditive inequality,

$$a(x, y) \leq a(x, z) + a(z, y) \qquad \text{for all } z,$$

and many interesting facts may be deduced from such inequalities and their ramifications. In particular, it follows by the subadditive ergodic theorem (see [53]) that the limiting velocity $\lim_{n\to\infty} a(0, nx)/n$ exists in every direction x.

One of the principal objects of study is the *wet region* at time t when liquid is supplied at the origin, i.e., the set $W(t) = \{x : a(0, x) \leq t\}$. The easiest way to describe the asymptotic behaviour of $W(t)$ for large t begins by 'filling in the holes'. Thus we define $\widetilde{W}(t) = W(t) + [-\frac{1}{2}, \frac{1}{2}]^d$, a region in $\mathbb{R}^d$. The set $\widetilde{W}(t)$ grows linearly as time passes, in the following sense. Subject to an appropriate moment condition on F, there exists a non-random set L having non-empty interior such that either

(a) L is compact and

$$(1-\epsilon)L \subseteq \frac{1}{t}\widetilde{W}(t) \subseteq (1+\epsilon)L \qquad \text{eventually, a.s.,}$$

for all $\epsilon > 0$, or

(b) $L = \mathbb{R}^d$, and

$$\frac{1}{t}\widetilde{W}(t) \supseteq [-M, M]^d \qquad \text{eventually, a.s.,}$$

for all $M > 0$.

Case (b) holds if and only if a typical time coordinate T satisfies $P(T = 0) \geq p_c$, which is to say that the set of edges with zero time coordinate forms a bond percolation process which is either *critical* or *supercritical*. The earliest such 'shape theorem' was proved by Richardson [65] in 1973. See [24, 51, 52] for more information and references.

Few non-trivial facts are known about the limit set L, and much effort has been spent, largely inconclusively, on attempting to decide whether L can ever be a Euclidean circle. More recently, interest has been concentrated on building a fluctuation theory for the set $\widetilde{W}(t)$: on what scale of t does $\widetilde{W}(t)$ differ from the dilated region tL? See [9] for example.

4. Epidemic models

It has long been realised that realistic models for the spread of disease must incorporate information about the interactions between individuals, and that 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 (see [45]).

The model is as follows. Let us suppose that individuals are placed at the vertices of the lattice $\mathbb{L}^d$ where $d \geq 1$, and let λ and δ be strictly positive constants. At each time t, the individual at x is in one 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 according to the following probabilistic rules. If the individual at x has state 1 (i.e., is ill) at time t, then it becomes susceptible during the short time interval $(t, t+h)$ with probability $\delta h + o(h)$. Here δ is the *rate of cure*. If the individual at x has state 2 (i.e., is susceptible) at time t, then it becomes ill during the interval $(t, t+h)$ with probability $\lambda n h + o(h)$ 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. (A full definition of the contact process involves a Markov process whose infinitesimal transition probabilities are given as above.)

A main question is whether or not the disease survives over all time intervals however long. Suppose that, at time 0, all individuals are susceptible except the origin which is ill. Does the disease spread, with strictly positive probability, throughout the entire infinity of space? More precisely, 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

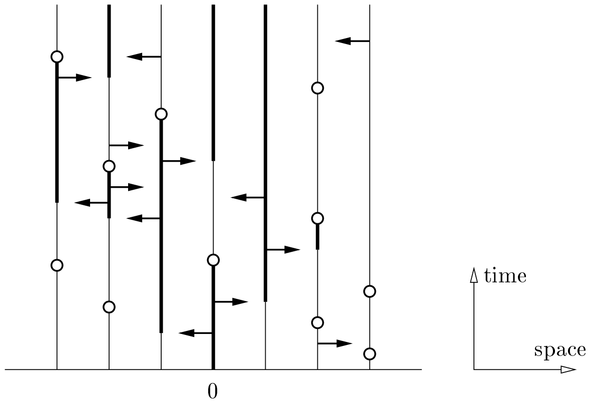


Fig. 4.1. The so called 'graphical representation' of the contact process when $d = 1$. The horizontal line represents 'space', and the vertical line above a point x represents the time axis at x. The marks $\circ$ are the points of cure, and the arrows are the arrows of infection. Suppose we are told that, at time 0, the origin is the unique infected point. Then all subsequent infections may be mapped 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 are infected.

so we may write $\psi(\lambda) = \psi(\lambda, 1)$. 'Evidently', $\psi(\lambda)$ is non-decreasing in λ, whence there exists a critical value λ_c such that

$$\psi(\lambda) \begin{cases} = 0 & \text{if } \lambda < \lambda_c \\ > 0 & \text{if } \lambda > \lambda_c. \end{cases}$$

The analogy with percolation is strong, with ψ taking the role of the percolation probability θ. Harris [45] proved amongst other things that λ_c is non-trivial, in the sense that $0 < \lambda_c < \infty$ if $d \geq 1$.

In Harris's first paper, he regarded the contact process as a Markov process ξ_t with a given structure. In a later paper [46], he exploited the fact that, suitably reformulated, the contact process *is* a percolation process of a certain type. His reformulation is as follows. We enrich the lattice $\mathbb{L}^d$ by considering the 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 with intensity δ; this requires in particular that a short interval $x \times (t, t+h)$ contains a point of cure with probability $\delta h + o(h)$ as $h \downarrow 0$. 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 4.1.

Here is the interpretation of these marks. Suppose that the vertex x is ill at time t. Then it remains infected until the first subsequent point of cure, i.e., until the time $T = \inf\{s > t : (x, s) \text{ is a point of cure}\}$. Meanwhile, whenever there exists an arrow oriented from x to y during the time-interval $x \times (t, T]$, then the infection at x spreads to y. (If y is already ill, the new infection has no effect.) It is a simple matter to check that the infection spreads in the manner of the contact process.

With infection originating at the origin only, then the infection continues forever if and only if S contains an infinite path which begins at the origin, moves in the direction of increasing time only, and is permitted to traverse infection arrows. See Figure 4.1 again. The probability of this event is nothing but the percolation probability for the oriented partly-continuous percolation system constructed above.

Once this link is made, it is not surprising that percolation technology may be adapted in order to study the contact process. Here is a major example of this. Two questions which remained open for some years were as follows.

- Is it the case that the critical contact process dies out, which is to say that $\psi(\lambda_c) = 0$?
- If $\lambda > \lambda_c$ and infection originating at the origin continues forever, then can one prove a 'shape theorem' for the manner of its spread? (Cf. the shape theorem of first-passage percolation.)

Building on the block arguments alluded to in Section 2.3, in 1991 Bezuidenhout and Grimmett [15] provided the final steps necessary to answer both questions affirmatively.

Slightly more realistic epidemic models require that individuals experience a period 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 showed in 1982 ([55]) that this system also is percolative, but in a different sense from that above. Let Δx denote the set of all neighbours of a vertex x. For each x, we draw oriented edges from x to some random subset of Δx 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 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 understand its geometry, thereby learning 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 very much harder to study, since they generally lack even the elementary property of monotonicity. In the 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 fail in more general systems for the following 'simple' reason. By adding an extra infective, one may subsequently infect a point which, during its removal period, prevents the infection from spreading further. Self-protection in a forest fire may be achieved by burning a pre-emptive firebreak.

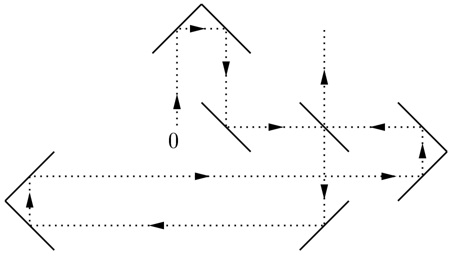


Fig. 5.1. A labyrinth of mirrors on the square lattice. The ray of light is reflected by the mirrors, and it is a problem is to determine, for a given density of mirrors, whether or not the light is a.s. restricted to a finite region.

See [25, 58] for further information about contact processes, and [13] for recent results concerning both the contact process (without recovery) and a more general contact process incorporating temporary removals.

5. Illumination of reflecting labyrinths

An electron travels through an environment of massive particles, suffering deflections when it impacts on these particles. In three essays [59] published in 1905, Hendrik Lorentz proposed a model sometimes referred to now as a 'Lorentz lattice gas'. Developed further by Ehrenfest under the name 'wind–tree model', and transferred to the square lattice, the physical phenomenon has given rise to a concrete problem of probability theory having substantial appeal.

Let $0 \leq p \leq 1$. We call each vertex x of the square lattice $\mathbb{L}^2$ a *mirror* with probability p, and a *crossing* otherwise; different vertices receive independent designations. Given that x is a mirror, we call it a *north-west (NW) mirror* with probability $\frac{1}{2}$ and a *north-east (NE) mirror* otherwise. We now place two-sided plane mirrors at vertices of $\mathbb{L}^2$ in the prescribed configuration (see Figure 5.1). A ray of light is shone northwards from the origin. When it strikes a crossing, it passes through undeflected. When it strikes a mirror, it is reflected through a right angle in the appropriate direction. It is not hard to see that: either the light traverses a semi-infinite path beginning at the origin (possibly with self-intersections) or it traverses a closed (finite) loop. Let $\xi(p)$ be the probability of the former situation, i.e., $\xi(p)$ is the probability that the light illuminates infinitely many vertices. The problem is to determine for which values of p (if any) it is the case that $\xi(p) > 0$.

In this lattice version of the Lorentz model, the mirrors represent the massive particles and the light represents the electron. There are continuum versions of the problem also (see [19, 69]).

It is apparently very difficult to determine whether or not $\xi(p) > 0$ for a given value of p. The only trivial fact is that $\xi(0) = 1$, and the only other known value is $\xi(1) = 0$. There are conflicting intuitions when $0 < p < 1$, and numerical simulations seem to suggest that $\xi(p) = 0$ whenever $p > 0$. The difficulty of the problem seems to lie in the fact that it is a mixture of a dynamical system and a random environment. Conditional on the environment of mirrors, the light behaves

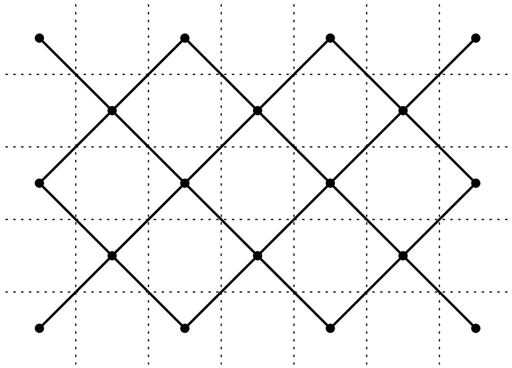


Fig. 5.2. When $p = 1$, the labyrinth of mirrors gives rise to a critical bond percolation process on a certain 'diagonal' copy $\mathcal{L}$ of the square lattice $\mathbb{L}^2$, drawn here in bold and broken lines respectively.

deterministically, but its trajectory can be very sensitive to minor changes in the positions of the mirrors.

That $\xi(1) = 0$ follows from a percolation argument. The history of this is slightly vague. It was certainly known in 1978, but appeared first in print in 1989 ([28]); further results appeared in [20]. The argument is as follows, and is illustrated in Figure 5.2. We work on an ancillary 'diagonal lattice' $\mathcal{L}$ having vertex set $(m+\frac{1}{2}, n+\frac{1}{2})$ for $m, n \in \mathbb{Z}$ with $m+n$ even; there is an edge joining $(m+\frac{1}{2}, n+\frac{1}{2})$ and $(r+\frac{1}{2}, s+\frac{1}{2})$ if and only if $|m-r| = |n-s| = 1$. We now use the mirrors of $\mathbb{L}^2$ to obtain a bond percolation process on $\mathcal{L}$. 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 the 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 process is a bond percolation model on a square lattice at density $\frac{1}{2}p$. When $p = 1$, this density equals $\frac{1}{2}$, and it is known that the percolation probability θ satisfies $\theta(\frac{1}{2}) = 0$ (see [44], or Section 1 of the current paper). This implies (by duality) that the origin of $\mathbb{L}^2$ is contained a.s. in the interior of some open circuit D of $\mathcal{L}$. Now, each edge of D corresponds to a superimposed mirror, and therefore D corresponds to a 'barrier' of mirrors surrounding the origin. Light cannot escape such a barrier, whence $\xi(1) = 0$.

There are numerous related systems which pose interesting challenges to the physicist and mathematician. For example, there are many other reflecting bodies than simple plane mirrors. In fact, in d (≥ 2) dimensions, there exist exactly

$$\sum_{s=0}^{d} \frac{(2d)!}{(2s)!2^{d-s}(d-s)!}$$

such bodies (see [31]). Very little indeed is known about the trajectories of light rays illuminating general 'random labyrinths' of mirrors.

In further work, Ruijgrok and Cohen [66] have proposed a study of 'rotator' models as well as 'mirror' models. In the simplest rotator model, a vertex of $\mathbb{L}^2$ is designated as one of three types: a 'right' rotator, a 'left' rotator, or a 'crossing'. As before, light traverses a crossing without deflection, but when incident on a right (resp. left) rotator, it is deflected 90° to the right (resp. left). It is not yet clear what methods may be used to develop a satisfactory mathematical theory of such systems.

In another development, one introduces a little extra randomness into the environment, as follows. Let $p_{\mathrm{rw}} > 0$. We designate each vertex a 'random-walk point' with probability p_{rw}; otherwise it may be a mirror or a crossing as above. When light is incident on a random-walk point, then a fair die is thrown in order to choose the exit direction; in dimension d, each of the $2d$ possible directions has equal probability $(2d)^{-1}$. Methods from percolation theory may be used in order to control the geometry of the ensuing labyrinth, and partial results follow. We state two of these briefly.

Consider a general reflecting labyrinth in two or more dimensions with a strictly positive density p_{rw} of random-walk points. If the density of non-trivial reflectors is sufficiently small (a reflector is called non-trivial if it is not the crossing) then

- the light illuminates an infinite set with strictly positive probability, and
- when this occurs, then the light is 'diffusive' in the sense that its position after n steps has (asymptotically) a normal (Gaussian) distribution with mean 0 and variance δn.

Here, δ is a strictly positive constant, which depends on the parameters of the labyrinth of mirrors. Further details may be found in [16, 31, 32, 34].

6. Ferromagnetism and random-cluster models

Perhaps the most famous spatial model of statistical physics is the Ising model for ferromagnetism. Founded in work of Lenz and Ising [48], this process has generated tremendous interest and has provided the setting for the development of a repertoire of techniques of wide applicability. The underlying physical phenomenon is the following. Consider the experiment of placing a piece of iron in a magnetic field; the field is increased from zero to some maximum, and then reduced back to zero. The iron may retain some residual magnetisation, but only when the temperature is not too high. There exists a critical temperature T_{c} marking the division between the two phases. In the Ising model, the iron is modelled by a part of a lattice, each vertex of which may be in either of two states labelled + and −. The states at neighbouring points interact in the manner of a so called 'Gibbs state' (see [27, 58]).

One generalisation of the Ising model is that proposed by Potts [64] in 1952. A feature of the Potts model is that each vertex may be in any of q distinct states, labelled $1, 2, \ldots, q$; the Ising model is recovered when $q = 2$. The Potts model also has a phase transition at some critical temperature $T_{\mathrm{c}}(q)$.

It is a remarkable fact that the Ising and Potts models, together with the percolation model, may be placed within a unified system having a coherent methodology. This was discovered in the late 1960s by Fortuin and Kasteleyn, and led to their formulation of a process which they called a 'random-cluster model'. Their construction is unusual, and the reader may wonder how Fortuin and Kasteleyn were led to it. The answer is that Kasteleyn observed that a number of different systems enjoy 'series and parallel laws'. The best known of these is electrical networks, in which two resistances of size r_1 and r_2 may be replaced by a single resistance of size r_1+r_2 if in series, or $(r_1^{-1}+r_2^{-1})^{-1}$ if in parallel. Percolation, Ising, and Potts models have a similar property, and Kasteleyn wished to understand whether these common properties indicated a more extensive common structure. A historical account may be found in [29].

The random-cluster model is as follows. Let $0 \le p \le 1$ and $q > 0$: these are the parameters of the system. Let $G = (V, E)$ be a fixed finite graph, and let F be a subset of E chosen according to the following probability function $\phi_{p,q}$:

$$\phi_{p,q}(F) = \frac{1}{Z}\, p^{|F|}(1-p)^{|E\setminus F|}q^{k(F)} \tag{6.1}$$

where $k(F)$ is the number of connected components of the graph (V, F), and Z is a constant which is chosen to ensure that $\sum_{F\subseteq E}\phi_{p,q}(F) = 1$.

Let us consider some special values. First, if $q = 1$, then

$$\phi_{p,1}(F) = p^{|F|}(1-p)^{|E\setminus F|},$$

which is to say that different edges are present independently of one another, each with probability p; this is bond percolation on G. If $q = 2, 3, \ldots$ then $\phi_{p,q}$ is related to the Potts model with q states, and in which p is a certain function of the temperature and 'pair-interaction'. In particular, $\phi_{p,2}$ corresponds to the Ising model. This, and many other useful facts, were discovered by Fortuin and Kasteleyn.

There is a difficulty which is not present for percolation, namely how to define a random-cluster model on an *infinite* graph; formula (6.1) simply does not work directly in this case. The answer, as with the Ising model, is to let $\phi_{\Lambda,p,q}$ be the random-cluster measure on a finite subgraph Λ of an infinite 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'. A full description would require a discussion of 'boundary conditions', and this is not appropriate here. Let $\phi_{p,q}$ denote the limiting probability measure for the whole lattice, and let

$$\theta(p,q) = \phi_{p,q}\big(0 \text{ belongs to an infinite cluster}\big)$$

be the corresponding percolation probability. It turns out that, for fixed $q \ge 1$, there exists a critical value of p, written $p_c(q)$, 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 $q = 2, 3, \ldots$ then $p_c(q)$ may be expressed in terms of the critical temperature $T_c(q)$ of

the q-state Potts model. In this sense (and indeed further) the phase transition of the random-cluster model generalises those of percolation, Ising, and Potts models.

Since the random-cluster model generalises so many systems of interest, it has been natural to develop a coherent theory thereof. A body of techniques has emerged in recent years, but many mysteries remain unresolved. For example, two known facts are:

- if $\theta(p,q) > 0$, there exists a.s. a *unique* infinite cluster,
- if q is large, say $q > Q(d)$, then $\theta(p_c(q), q) > 0$.

[See [30, 31, 54, 56]. Actually some technical assumptions involving boundary conditions are needed for these conclusions.] The second fact is particularly striking, since it implies that $\theta(\cdot, q)$ is *discontinuous* at the critical point, in contrast to the corresponding conjecture for percolation (i.e., when $q = 1$).

In contrast, one may conjecture that

- if q is small, say $1 \le q < Q(d)$, then $\theta(p_c(q), q) = 0$,
- if $q \ge 1$ and $p < p_c(q)$, then $\phi_{p,q}\big(\mathrm{rad}(C) \ge n\big)$ decays exponentially as $n \to \infty$ (cf. (2.2)).

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 $1 \le q < 4$, to the belief that $Q(2) = 4$ and $Q(d) = 2$ when $d \ge 6$. In addition, very little is known when $0 < q < 1$.

This beautiful model is an outstanding challenge to mathematicians. It promises a unified structure which will explain further the Ising and Potts models, and which places them 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 [6, 12, 30, 31, 35, 36].

Acknowledgements. This essay was written in Voulangis, following work supported in part by the European Union under contracts CHRX–CT93–0411 and FMRX–CT96–0075A and by the Engineering and Physical Sciences Research Council of the UK under grant GR/L15426.

References

1. Aharony, A. and Stauffer, D., *Introduction to Percolation Theory* (Second edition), Taylor and Francis, 1991.
2. Aizenman, M., *Geometric analysis of ϕ^4 fields and Ising models*, Communications in Mathematical Physics **86** (1982), 1–48.
3. Aizenman, M., *The geometry of critical percolation and conformal invariance*, Proceedings STATPHYS 1995 (Xianmen) (Hao Bai-lin, ed.), World Scientific, 1995.
4. Aizenman, M. and Barsky, D. J., *Sharpness of the phase transition in percolation models*, Communications in Mathematical Physics **108** (1987), 489–526.
5. Aizenman, M., Barsky, D. J., and Fernández, R., *The phase transition in a general class of Ising-type models is sharp*, Journal of Statistical Physics **47** (1987), 343–374.
6. Aizenman, M., Chayes, J. T., Chayes, L., and Newman, C. M., *Discontinuity of the magnetization in one-dimensional $1/|x-y|^2$ Ising and Potts models*, Journal of Statistical Physics **50** (1988), 1–40.
7. Aizenman, M., Kesten, H., and Newman, C. M., *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-Verlag, Berlin, 1987, pp. 13–20.
8. Aizenman, M. and Newman, C. M., *Tree graph inequalities and critical behavior in percolation models*, Journal of Statistical Physics **36** (1984), 107–143.
9. Alexander, K., *Approximation of subadditive functions and convergence rates in limiting-shape results*, Annals of Probability **25** (1997), 30–55.
10. Alexander, K. S., Chayes, J. T., and Chayes, L., *The Wulff construction and asymptotics of the finite cluster distribution for two-dimensional Bernoulli percolation*, Communications in Mathematical Physics **131** (1990), 1–50.
11. Barsky, D. J., Grimmett, G. R., and Newman, C. M., *Percolation in half spaces: equality of critical probabilities and continuity of the percolation probability*, Probability Theory and Related Fields **90** (1991), 111–148.
12. Baxter, R. J., *Exactly Solved Models in Statistical Mechanics*, Academic Press, London, 1982.
13. Berg, J. van den, Grimmett, G. R., and Schinazi, R. B., *Dependent random graphs and spatial epidemics*, Annals of Applied Probability (1997) (to appear).
14. Berg, J. van den and Keane, M., *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., 1984, pp. 61–65.
15. Bezuidenhout, C. E. and Grimmett, G. R., *The critical contact process dies out*, Annals of Probability **18** (1990), 1462–1482.
16. Bezuidenhout, C. E. and Grimmett, G. R., *A central limit theorem for random walks in random labyrinths* (1997) (to appear).
17. Broadbent, S. R., *Discussion on symposium on Monte Carlo methods*, Journal of the Royal Statistical Society B **16** (1954), 68.
18. Broadbent, S. R. and Hammersley, J. M., *Percolation processes I. Crystals and mazes*, Proceedings of the Cambridge Philosophical Society **53** (1957), 629–641.
19. Bunimovich, L. A. and Sinai, Ya. G., *Statistical properties of Lorentz gas with periodic configuration of scatterers*, Communications in Mathematical Physics **78** (1981), 479–497.
20. Bunimovitch, L. A. and Troubetzkoy, S. E., *Recurrence properties of Lorentz lattice gas cellular automata*, Journal of Statistical Physics **67** (1992), 289–302.
21. Burton, R. M. and Keane, M., *Density and uniqueness in percolation*, Communications in Mathematical Physics **121** (1989), 501–505.
22. Cardy, J., *Critical percolation in finite geometries*, Journal of Physics A: Mathematical and General **25** (1992), L201.
23. Chayes, J. T., Chayes, L., Grimmett, G. R., Kesten, H., and Schonmann, R. H., *The correlation length for the high density phase of Bernoulli percolation*, Annals of Probability **17** (1989), 1277–1302.
24. Cox, J. T. and Durrett, R. T., *Some limit theorems for percolation processes with necessary and sufficient conditions*, Annals of Probability **9** (1981), 583–603.
25. Durrett, R. T., *The contact process, 1974–1989*, Mathematics of Random Media, Blacksburg, VA, Lectures in Applied Mathematics, vol. 27, AMS, 1991, pp. 1–18.
26. Efros, A. L., *Physics and Geometry of Disorder*, Mir, Moscow, 1986.
27. Georgii, H.-O., *Gibbs Measures and Phase Transitions*, Walter de Gruyter, Berlin, 1988.
28. Grimmett, G. R., *Percolation*, Springer-Verlag, Berlin, 1989.
29. Grimmett, G. R., *The random-cluster model*, Probability, Statistics and Optimisation (F. P. Kelly, ed.), John Wiley & Sons, Chichester, 1994, pp. 49–63.
30. Grimmett, G. R., *The stochastic random-cluster process and the uniqueness of random-cluster measures*, Annals of Probability **23** (1995), 1461–1510.
31. Grimmett, G. R., *Percolation and disordered systems*, Ecole d'Eté de Probabilités de Saint Flour XXVI–1996 (P. Bernard, ed.), Lecture Notes in Mathematics, Springer-Verlag, Berlin, 1997, pp. 153–300 (to appear).
32. Grimmett, G. R., *Stochastic pin-ball* (1997) (to appear).
33. Grimmett, G. R. and Marstrand, J. M., *The supercritical phase of percolation is well behaved*, Proceedings of the Royal Society (London), Series A **430** (1990), 439–457.
34. Grimmett, G. R., Menshikov, M. V., and Volkov, S. E., *Random walks in random labyrinths*, Markov Processes and Related Fields **2** (1996), 69–86.

35. Grimmett, G. R. and Piza, M. S. T., *Decay of correlations in subcritical Potts and random-cluster models*, Communications in Mathematical Physics (1996) (to appear).
36. Häggström, O., *Random-cluster representations in the study of phase transitions* (1996) (to appear).
37. Hammersley, J. M., *Percolation processes. Lower bounds for the critical probability*, Annals of Mathematical Statistics **28** (1957), 790–795.
38. Hammersley, J. M., *Bornes supérieures de la probabilité critique dans un processus de filtration*, Le Calcul de Probabilités et ses Applications, CNRS, Paris, 1959, pp. 17–37.
39. Hammersley, J. M., *Origins of percolation theory*, Annals of the Israel Physical Society **5** (1983), 47–57.
40. Hammersley, J. M. and Morton, K. W., *Poor man's Monte Carlo*, Journal of the Royal Statistical Society B **16** (1954), 23–38.
41. Hammersley, J. M. and Welsh, D. J. A., *First-passage percolation, subadditive processes, stochastic networks and generalized renewal theory*, Bernoulli, Bayes, Laplace Anniversary Volume (J. Neyman and L. LeCam, eds.), Springer-Verlag, Berlin, 1965, pp. 61–110.
42. Hara, T. and Slade, G., *Mean-field critical behaviour for percolation in high dimensions*, Communications in Mathematical Physics **128** (1990), 333–391.
43. Hara, T. and Slade, G., *Mean-field behaviour and the lace expansion*, Probability and Phase Transition (G. R. Grimmett, ed.), Kluwer, Dordrecht, 1994, pp. 87–122.
44. Harris, T. E., *A lower bound for the critical probability in a certain percolation process*, Proceedings of the Cambridge Philosophical Society **56** (1960), 13–20.
45. Harris, T. E., *Contact interactions on a lattice*, Annals of Probability **2** (1974), 969–988.
46. Harris, T. E., *Additive set-valued Markov processes and graphical methods*, Annals of Probability **6** (1978), 355–378.
47. Hughes, B. D., *Random Walks and Random Environments. Vol. 2. Random Environments*, Oxford University Press, New York, 1996.
48. Ising, E., *Beitrag zur Theorie des Ferromagnetismus*, Zeitschrift für Physik **31** (1925), 253–258.
49. Kesten, H., *The critical probability of bond percolation on the square lattice equals* $\frac{1}{2}$, Communications in Mathematical Physics **74** (1980), 41–59.
50. Kesten, H., *Percolation Theory for Mathematicians*, Birkhäuser, Boston, 1982.
51. Kesten, H., *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-Verlag, Berlin, 1986, pp. 125–264.
52. Kesten, H., *Percolation theory and first-passage percolation*, Annals of Probability **15** (1987), 1231–1271.
53. Kingman, J. F. C., *Subadditive processes*, Ecole d'Eté de Probabilités de Saint Flour V–1975 (P. L. Hennequin, ed.), Lecture Notes in Mathematics no. 539, Springer-Verlag, Berlin, 1976, pp. 167–223.
54. Kotecký, R. and Shlosman, S., *First order phase transitions in large entropy lattice systems*, Communications in Mathematical Physics **83** (1982), 493–515.
55. Kuulasmaa, K., *The spatial general epidemic and locally dependent random graphs*, Journal of Applied Probability **19** (1982), 745–758.
56. Laanait, L., Messager, A., Miracle-Sole, S., Ruiz, J., and Shlosman, S., *Interfaces in the Potts model I: Pirogov–Sinai theory of the Fortuin–Kasteleyn representation*, Communications in Mathematical Physics **140** (1991), 81–91.
57. Langlands, R., Pouliot, P., and Saint-Aubin, Y., *Conformal invariance in two-dimensional percolation*, Bulletin of the American Mathematical Society **30** (1994), 1–61.
58. Liggett, T. M., *Interacting Particle Systems*, Springer-Verlag, Berlin, 1985.
59. Lorentz, H. A., *The motion of electrons in metallic bodies, I, II, and III*, Koninklijke Akademie van Wetenschappen te Amsterdam, Section of Sciences **7** (1905), 438–453, 585–593, 684–691.
60. Madras, N. and Slade, G., *The Self-Avoiding Walk*, Birkhäuser, Boston, 1993.
61. Menshikov, M. V., *Coincidence of critical points in percolation problems*, Soviet Mathematics Doklady **33** (1986), 856–859.

62. Menshikov, M. V., Molchanov, S. A., and Sidorenko, A. F., *Percolation theory and some applications*, Itogi Nauki i Techniki (Series of Probability Theory, Mathematical Statistics, Theoretical Cybernetics) **24** (1986), 53–110.
63. Peierls, R., *On Ising's model of ferromagnetism*, Proceedings of the Cambridge Philosophical Society **36** (1936), 477–481.
64. Potts, R. B., *Some generalized order-disorder transformations*, Proceedings of the Cambridge Philosophical Society **48** (1952), 106–109.
65. Richardson, D., *Random growth in a tesselation*, Proceedings of the Cambridge Philosophical Society **74** (1973), 515–528.
66. Ruijgrok, T. W. and Cohen, E. G. D., *Deterministic lattice gas models*, Physics Letters A **133** (1988), 415–418.
67. Russo, L., *A note on percolation*, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete **43** (1978), 39–48.
68. Seymour, P. D. and Welsh, D. J. A., *Percolation probabilities on the square lattice*, Advances in Graph Theory (B. Bollobás, ed.), Annals of Discrete Mathematics 3, North-Holland, Amsterdam, 1978, pp. 227–245.
69. Spohn, H., *Large Scale Dynamics of Interacting Particles*, Springer-Verlag, Berlin, 1991.
70. Sykes, M. F. and Essam, J. W., *Exact critical percolation probabilities for bond and site problems in two dimensions*, Physical Review Letters **10** (1963), 3–4.
71. Sykes, M. F. and Essam, J. W., *Exact critical percolation probabilities for site and bond problems in two dimensions*, Journal of Mathematical Physics **5** (1964), 1117–1127.
72. Wierman, J. C., *Bond percolation on honeycomb and triangular lattices*, Advances in Applied Probability **13** (1981), 293–313.

Statistical Laboratory, University of Cambridge, 16 Mill Lane, Cambridge CB2 1SB, United Kingdom
E-mail address: g.r.grimmett@statslab.cam.ac.uk
URL: http://www.statslab.cam.ac.uk/~grg/