1
Introduction

1.1 Motivation for continuum models

Many phenomena in physics, chemistry and biology can be modelled by spatial random processes where the randomness is in the geometry of the space rather than in the random behaviour or motion of an object in a deterministic setting. As typical examples of the phenomena we have in mind, consider the spread of a disease in an orchard where the trees are arranged in a grid, and where the disease spreads from an infected tree to its neighbouring trees. In this example, the owner of the orchard is interested in the probability that a particular disease will eventually kill all the trees in the orchard. Another example is the process of the ground getting wet during a period of rain. The randomness here is the place where the raindrops fall on the ground and the size of the wetted region per raindrop. Finally, consider the spread of a disease in a forest. The infection is transmitted from one tree to another, which need not be in the vicinity of the infected tree. This is more likely to happen when the trees are closely spaced than when they are far apart. The collection of infected trees forms a random subset of trees in the forest.

The geometric structure of the first example is discrete, whereas in the next two examples, although the number of raindrops or trees is countable, the position of either is in the continuous space. A rigorous mathematical model to describe the first example is the standard discrete percolation model. This model has been studied extensively in the last three decades and an excellent reference on the mathematical aspects of this model is the book by Grimmett (1989). The second and third examples are usually described by a continuum percolation model; such models are the subject of this book.

Some geometric aspects of the continuum percolation model have been studied in the context of stochastic geometry. In the language of stochastic geometry, the continuum percolation model is usually referred to as a coverage process or a Boolean model.

To get an idea of the kind of questions addressed in a percolation-theoretic study, we elaborate the examples of rainfall and the spread of a disease in a forest. Before the rain starts, the ground is assumed to be completely dry. At a point where a raindrop falls, the ground soaks up the water and a circular wet patch is formed. When the first raindrops fall, we see small wet regions inside a large dry region. The wet region grows when more raindrops reach the ground and at some instant, so many raindrops have reached the ground that the picture suddenly changes from wet 'islands' inside a large dry region to dry 'islands' inside a large wet region. This phenomenon of a sudden drastic change in the global spatial structure is called a *phase transition*. Typically, the parameter of the model is not the time, but the density of raindrops on the ground. So for instance, we say that the phase transition takes place at a given density of the raindrops, rather than at a given time. The nature of such phase transitions is an important subject in the percolation-theoretic study of Boolean models. In the example of the spread of an infection in a forest, a question of interest is whether the infection of one particular tree may result in the infection being transmitted to a tree far away. This is of course more likely when the density of the trees in the forest is high. Based on the density of the trees, a phase transition formulation may be obtained for this model too.

The focus of this book is on mathematically rigorous results in models of continuum percolation. In this context, we remark that there are many results available in the applied literature which have yet to be mathematically verified.

1.2 Discrete percolation

Before we introduce continuum percolation models we present a short treatment of discrete models. There are several reasons for doing this. First, independent percolation on the integer lattice (to be defined below) was the first percolation model studied and many of the ideas in the theory of this model can be used in the study of continuum models as well; many of the results in the continuum are analogues of discrete results. Secondly, discrete percolation models are in some sense the simplest percolation models to describe and they are suitable for the reader to get a feeling for the types of problems which are involved. Finally, an important technique in the theory of continuum percolation is to approximate the continuum model by a discrete one. In these instances, we need to know something about discrete percolation. Our treatment of discrete percolation is concise, and we refer the reader to Grimmett (1989) for a detailed discussion of discrete percolation. For proofs which are not given here we refer to either Grimmett (1989) or the references in the notes.

The setup is as follows. Each element of the d-dimensional integer lattice $\mathbf{Z}^d$ is a vertex, where $d \geq 1$. Two vertices at a Euclidean distance one apart are called *neighbours*. Each pair of neighbours has an edge between them. The graph obtained this way is, with a slight abuse of notation, denoted by $\mathbf{Z}^d$. An edge is often called a *bond*, and the set of all bonds is denoted by $\mathbf{E}$. Bonds can be either *open* or *closed*. A *path* is a finite or infinite alternating sequence $(z_1, e_1, z_2, e_2, \ldots)$ of vertices z_i and bonds e_i such that $z_i \neq z_j$ and $e_i \neq e_j$ whenever $i \neq j$ and such that e_i is the bond between the neighbours z_i and z_{i+1}, for all i. The length of a path is the number of bonds it contains. A *circuit* is a finite path, the only difference being that it starts and ends at the same vertex. An *open (closed) path* is a path whose bonds are all open (closed). Two vertices are said to be *connected* if there is a finite open path from one to the other. An open *cluster* is a set of connected vertices which is maximal with respect to this property. Of course, clusters can be either finite or infinite. The open cluster containing the origin is denoted by $C(0)$.

Next we introduce probability. For $0 \leq p \leq 1$, we equip the space $\Omega = \{0, 1\}^{\mathbf{E}}$ with the natural product measure P_p, which is defined via $P_p(\omega(e) = 1) = p$ for all $e \in \mathbf{E}$. For any realisation $\omega \in \Omega$, the bond e is said to be open if $\omega(e) = 1$ and closed otherwise. Thus each bond is open with probability p independently of all other bonds.

This is the basic percolation model on the d-dimensional integer lattice. We are interested in unbounded clusters, so here are some natural definitions:

Definition 1.1 *The percolation function $\theta^{(d)}$ is defined by*

$$\theta^{(d)}(p) = P_p(\text{card}(C(0)) = \infty).$$

We define the function $\chi^{(d)}(p)$ by

$$\chi^{(d)}(p) = E_p(\text{card}(C(0))),$$

where E_p is the expectation operator corresponding to P_p, and card$(\cdot)$ *denotes cardinality.*

Much of the theory of discrete percolation is concerned with the behaviour of these functions. It seems obvious that both $\theta^{(d)}$ and $\chi^{(d)}$ are non-decreasing in p. In Chapter 2 it will become clear how to prove this. Based on $\theta^{(d)}$ and $\chi^{(d)}$ we can define the following *critical probabilities*:

Definition 1.2 *The critical probability $p_c(d)$ is defined by*

$$p_c(d) = \inf\{p : \theta^{(d)}(p) > 0\}.$$

The critical probability $p_T(d)$ is defined by

$$p_T(d) = \inf\{p : \chi^{(d)}(p) = \infty\}.$$

The two critical probabilities just defined are quite natural. There is a third critical probability which may not seem that natural at first sight but which turns out to be very useful. To define this critical probability, let $\sigma_p((n_1, n_2, \ldots, n_d), i)$ be the probability that the box $[0, n_1] \times [0, n_2] \times \cdots \times [0, n_d]$ contains an open path connecting two opposite faces in the i-th direction.

Definition 1.3 *For $d \geq 2$, the critical probability $p_S(d)$ is defined by*

$$p_S(d) = \inf\{p : \limsup_{n\to\infty} \sigma_p((n, 3n, 3n, \ldots, 3n), 1) = 0\}.$$

It is obvious that $p_T(d) \leq p_c(d)$ for all d. It is also clear that $p_c(1) = p_T(1) = 1$. Other properties of these critical probabilities are not so easy to obtain:

Theorem 1.1 *For all $d \geq 2$ we have $0 < p_c(d) < 1$.*

Theorem 1.2 *For all $d \geq 2$ we have $p_c(d) = p_T(d) = p_S(d)$.*

Actual values are known only in one and two dimensions. It is obvious that $p_c(1) = 1$ and we also know that $p_c(2) = \frac{1}{2}$. This last result is far from trivial! The proof of Theorem 1.2 is very hard and we do not give it here. Theorem 1.1 lies at the very heart of percolation theory. It establishes the existence of a *phase transition*; i.e. the macroscopic behaviour of the system is very different for values of p below and above the critical probability $p_c(d)$. If an infinite cluster exists we say that *percolation* occurs. The idea behind the proof of Theorem 1.1 will be used a few more times in this book, so we present the proof here:

Proof of Theorem 1.1 The inequality $p_c(d) > 0$ is very simple. Indeed, the number of distinct paths of length n starting at the origin is at most $2d(2d-1)^{n-1}$. (For the first bond, we have $2d$ possibilities; after that we have at most $2d-1$ possibilities for each new bond because we are not allowed to go back to where we came from.) Each of these paths has probability p^n to be open. Thus the expected number of open paths of length n starting at the origin is at most $2d(2d-1)^{n-1}p^n$. If $p < (2d-1)^{-1}$ then $\sum_{n=1}^{\infty} 2d(2d-1)^{n-1}p^n < \infty$, and hence the expected number of open edges in the component $C(0)$ is finite. This necessarily means that the probability that the component $C(0)$ is finite is equal to 1 and hence $\theta^{(d)}(p) = 0$ if $p < (2d-1)^{-1}$. Thus we obtain that $p_c(d) \geq (2d-1)^{-1}$.

For the other inequality we observe that it suffices to prove it for the case $d = 2$ as $p_c(d)$ is clearly non-increasing in d. We need to introduce the *dual* graph $\mathbf{Z}^{2*}$. This is the graph obtained from $\mathbf{Z}^2$ by shifting it by the vector $(\frac{1}{2}, \frac{1}{2})$. The set of edges of the dual graph is denoted by $\mathbf{E}^*$. Each edge of $\mathbf{E}$ now crosses exactly one edge in $\mathbf{E}^*$. We declare an edge in $\mathbf{E}^*$ to be open if and only if the edge it crosses in $\mathbf{E}$ is open, and closed otherwise. It is intuitively obvious and a well-known fact in graph theory (Whitney 1933) that there is a closed circuit in $\mathbf{Z}^{2*}$ surrounding the origin if and only if $C(0)$ is finite. Now we can perform a counting argument as in the first part of this proof. There are at most $n3^n$ distinct circuits of length n surrounding the origin. (This is a rather crude bound: such a circuit has to contain at least one vertex on the x-axis. There are at most n possibilities for this. Starting at this vertex, we have only three possibilities for each new bond.) If for some $N > 0$ (i) all bonds in $[-N, N] \times [-N, N]$ are open and (ii) there is no closed circuit in the dual surrounding $[-N, N] \times [-N, N]$, then $C(0)$ is infinite. The event in (i) certainly has positive probability. Furthermore, a circuit surrounding $[-N, N] \times [-N, N]$ has length at least $4N$. Hence, if $p > \frac{2}{3}$ we can choose N so large that $\sum_{n=4N}^{\infty} n3^n(1-p)^n < 1$ and for such p and N, the event in (ii) also has positive probability. Because the events in (i) and (ii) depend on disjoint sets of edges, they are independent, and we conclude that $p_c(2) \leq \frac{2}{3}$. □

Now that we have established the existence of infinite open clusters for $p > p_c$, the question arises of just how many infinite open clusters exist. There is a remarkable answer to that question. First observe that the existence of an infinite open cluster does not depend on the state of any finite set of bonds. Hence it follows from Kolmogorov's $0-1$ law that the existence of an infinite open cluster has probability either zero or one. This, of course, corresponds to the different phases of the percolation model: for $p < p_c(d)$ there is no infinite open cluster a.s. and for $p > p_c(d)$, the probability of having an infinite open cluster is positive and hence equal to 1. What happens *at* the critical probability is known in two dimensions and in dimension higher than 19 only: there is no infinite cluster a.s. in these cases. The remarkable fact referred to above is the following:

Theorem 1.3 *There is at most one infinite open cluster a.s.*

This result is referred to as the *uniqueness of the infinite cluster*.

We continue the discussion with some basic inequalities which are very useful in the analysis of models of this type. For this, we need to introduce some terminology. There is a natural partial order on $\Omega = \{0, 1\}^{\mathbf{E}}$: $\omega \leq \omega'$ if and only if $\omega(e) \leq \omega'(e)$ for all $e \in \mathbf{E}$. An event A in Ω (we assume that Ω is equipped with the usual Borel σ-field) is said to be *increasing* if its indicator function is increasing (a real-valued function f on Ω is increasing if $f(\omega) \leq f(\omega')$ whenever $\omega \leq \omega'$). An event A is said to be *decreasing* if its complement is increasing. A typical example of an increasing event is the event that two distinct vertices are connected to each other by an open path.

More generally, we can consider a product space $\Omega_k = \{0, 1, \ldots, k\}^{\Sigma}$, where Σ is a countable set, and equip Ω_k with product measure $P_p = \{p_0, p_1, \ldots, p_k\}^{\Sigma}$, where $\sum_{i=0}^{k} p_i = 1$. There is a natural partial order on Ω_k and the notions of increasing and decreasing events generalise easily. Writing P_p and E_p for probabilities and expectations with respect to p, we have the following important inequality:

Theorem 1.4 (FKG inequality) *Let f_1 and f_2 be both increasing or both decreasing functions. Then*

$$E_p f_1 f_2 \geq E_p f_1 E_p f_2.$$

Taking f_1 and f_2 to be the indicator functions of two increasing (or two decreasing) events A and B, respectively, this inequality reduces to

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

This result is not surprising: if there exists an open path connecting two different vertices, another path connecting two other vertices becomes more likely as it can 'use' the bonds of the first path.

Sometimes we need an inequality which goes in the opposite direction. Given the existence of an open path connecting two vertices, we can make it 'harder' for other connections to exist by requiring them to be disjoint from the first connection. This motivates the following definition. Suppose A and B are increasing events which depend only on the state of finitely many bonds. We define $A \square B$ to be the set of all configurations ω for which there exist disjoint sets of open bonds with the property that the first such set guarantees the occurrence of A and the second guarantees the occurrence of B. More precisely, $A \square B$ is the set of all configurations ω for which there exist finite and disjoint sets of bonds K_A and K_B such that any configuration ω' with $\omega'(e) = 1$, for all $e \in K_A$, is in A, and any configuration ω'' with $\omega''(e) = 1$, for all $e \in K_B$, is in B.

Theorem 1.5 (BK inequality) *Let A and B be two increasing events which depend on the state of only finitely many bonds. Then*

$$P_p(A \square B) \leq P_p(A)P_p(B).$$

The requirement that A and B are allowed to depend on only finitely many bonds has a technical reason. As we shall see, this will not be important in applications. In the next chapter, we shall derive continuum analogues of Theorems 1.4 and 1.5.

Next, we discuss a method for estimating the rate of change of $P_p(A)$ as a function of p, for increasing events A. Again, we assume that A depends on the state of finitely many edges only. For this, we need yet another definition. A bond e is said to be *pivotal* for A if $1_A(\omega) \neq 1_A(\omega_e)$, where ω_e is the configuration obtained from ω by changing the value at e; i.e. $\omega_e(e) = 1 - \omega(e)$. In words, a bond is pivotal for A if the occurrence or non-occurrence of A depends crucially on the state of the bond e. It is intuitively clear that the rate of change of $P_p(A)$ as a function of p is related to the number of pivotal bonds.

Theorem 1.6 (Russo's formula) *Let A be an increasing event which depends on only finitely many bonds. Then*

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

In the discussion so far, we assumed that bonds were either open or closed with certain probabilities. There was no randomness in the vertices at all. This is the reason that we call this model *bond percolation*. But we could as well declare vertices instead of edges to be open or closed with probability p and $1 - p$ respectively, obtaining a *site* model. The discussion of this site model is similar to the discussion of the bond model above. All results in this section have a natural analogue in the site setting, although the value of the critical probability for independent site percolation on the two-dimensional integer lattice is not known. We shall use these results in the site setting freely with a possible reference to the result in the bond setting.

As mentioned before, discretisation is an important technique in the theory of continuum percolation. Sometimes we end up with a more complicated discrete lattice structure than the nearest-neighbour integer lattice. Also it might be the case that there are different types of sites which are open with different probabilities. Take the d-dimensional integer lattice and draw an edge between any two vertices v and w for which $|v - w| \leq 2L$, where L is some positive constant. The graph obtained this way is denoted by $\mathcal{G}_L$. We can perform independent site percolation on this new graph, and this leads to the critical values $p_c(\mathcal{G}_L)$, $p_T(\mathcal{G}_L)$ and $p_S(\mathcal{G}_L)$.

This site-percolation model can also be extended to a multi-parametric setting. For example, consider a 'two-layered graph' $\mathcal{G}_{(L_1,L_2)}$ which is defined as follows. We place a copy of $\mathcal{G}_{L_2}$ 'above' $\mathcal{G}_{L_1}$ and we draw an edge between $v \in \mathcal{G}_{L_1}$ and $w \in \mathcal{G}_{L_2}$ if $d(v, w) \leq L_1 + L_2$ (here we abuse notation: v and w are viewed as elements of $\mathbb{R}^d$). Thus a vertex $v \in \mathcal{G}_{L_1}$ and a vertex $w \in \mathcal{G}_{L_2}$ are adjacent if and only if $d(v, w) \leq L_1 + L_2$. Now we perform multi-parameter independent site percolation on $\mathcal{G}_{(L_1,L_2)}$ by declaring a site in $\mathcal{G}_{L_1}$ to be open with probability p_1 and a site in $\mathcal{G}_{L_2}$ to be open with probability p_2. Rather than a critical point, in this model we can define a region inside the unit square where percolation occurs:

$$p_c(\mathcal{G}_{(L_1,L_2)}) = \{(p_1, p_2) : P_{(p_1,p_2)}(C(0) = \infty) > 0\},$$

where $C(0)$ is the union of the open clusters of the origins in $\mathcal{G}_{L_1}$ and $\mathcal{G}_{L_2}$. The regions $p_T(\mathcal{G}_{(L_1,L_2)})$ and $p_S(\mathcal{G}_{(L_1,L_2)})$ are defined similarly. The result which we shall need is a generalisation of Theorem 1.2:

Theorem 1.7 *In the setting just described it is the case that* $p_c(\mathcal{G}_L) = p_T(\mathcal{G}_L) = p_S(\mathcal{G}_L)$, *and* $p_c(\mathcal{G}_{(L_1,L_2)}) = p_T(\mathcal{G}_{(L_1,L_2)}) = p_S(\mathcal{G}_{(L_1,L_2)})$.

We end this section with a short discussion on mixed bond/site models. In such a model, both the sites and bonds of the integer lattice are either open or closed with certain probabilities. In its most general form, we have a parameter p and for each bond or vertex, w say, there is a non-decreasing function f_w such that w is open with probability $f_w(p)$, independently of all other bonds and vertices. Many of the results quoted thus far have their analogues in the mixed setting. In particular, Theorem 1.2 is still true in this setting. Here is a version of Russo's formula for this particular setting which we shall need in Chapter 6:

Theorem 1.8 (Russo's formula) *Consider a mixed bond/site model and let* A *be an increasing event which depends on the state of only finitely many vertices and bonds. Suppose in addition that there are non-decreasing differentiable functions* f_b, *such that the bond or vertex* b *is open with probability* $f_b(p)$ *independently of all other vertices and bonds. Then*

$$\frac{d}{dp}P_p(A) = \sum_{e \in \mathbf{E}} P_p(e \text{ is pivotal for } A)\frac{d}{dp}f_e(p) + \sum_{v \in \mathbf{Z}^d} P_p(v \text{ is pivotal for } A)\frac{d}{dp}f_v(p).$$

1.3 Stationary point processes

In the discrete percolation model of the previous section, the vertices of the random graph under consideration were non-random; they were formed by the elements of the d-dimensional integer lattice. In models for continuum percolation, this is no longer the case. The positions of the vertices themselves are random, and they are formed by the occurrences of a *stationary point process*. In this section, we introduce point processes, derive some basic properties and give some examples.

One can think of a point process as a random set of points in space. But of course, this is not a very mathematical definition, and we have to make precise what we mean by 'random' here. A natural way to do this is the following. Denote the σ-algebra of Borel sets in $\mathbb{R}^d$ by $\mathcal{B}^d$, and denote by N the set of all counting measures on $\mathcal{B}^d$ which assign finite measure to bounded Borel sets and for which the measure of a point is at most 1. In this way, N can be identified with the set of all configurations of points in $\mathbb{R}^d$ without limit points. We equip N with the σ-algebra $\mathcal{N}$ generated by sets of the form

$$\{n \in N : n(A) = k\},$$

where $A \in \mathcal{B}^d$ and k is an integer. A point process can now be defined as follows:

Definition 1.4 A point process X *is a measurable mapping from a probability space* $(\Omega, \mathcal{F}, P)$ *into* $(N, \mathcal{N})$.

The *distribution* of X is the measure μ on $\mathcal{N}$ induced by X; i.e. μ is defined through the equation $\mu(G) = P(X^{-1}(G))$, for all $G \in \mathcal{N}$. The definition of $\mathcal{N}$ allows us to count the number of points in a set $A \in \mathcal{B}^d$: the mapping $f_A : N \to N$ defined by $f_A(n) = n(A)$ is measurable by the very construction of $\mathcal{N}$. Hence the composition $f_A \circ X : \Omega \to N$ is a random variable which we denote by $X(A)$. In words, $X(A)$ represents the random number of points inside A.

In continuum models, we do not have a nice periodic structure as in discrete percolation models. The requirement that the lattice in discrete percolation is periodic is replaced by the requirement that the point process X is *stationary*. Let T_t be the translation in $\mathbb{R}^d$ over the vector t: $T_t(s) = t + s$, for all $s \in \mathbb{R}^d$. Then T_t induces a transformation $S_t : N \to N$ through the equation

$$(S_t n)(A) = n(T_t^{-1}(A)), \tag{1.1}$$

for all $A \in \mathcal{B}^d$. On a higher level, S_t induces a transformation $\tilde{S}_t$ on measures

μ on $\mathcal{N}$ through the equation

$$(\tilde{S}_t\mu)(G) = \mu(S_t^{-1}G), \tag{1.2}$$

for all $G \in \mathcal{N}$. Now we can define stationarity:

Definition 1.5 *The point process X is said to be* stationary *if its distribution is invariant under $\tilde{S}_t$ for all $t \in \mathbb{R}^d$.*

Definition 1.6 *The finite-dimensional (fidi) distributions of a point process X are the joint distributions, for all finite families of bounded Borel sets $A_1, \ldots, A_k$, of the random variables $X(A_1), \ldots, X(A_k)$.*

Standard methods (see e.g. Daley and Vere-Jones 1988) show that the distribution of a point process X is completely determined by its fidi distributions. The fidi distributions are thus one way of specifying a point process. In Chapter 7, we shall introduce a completely different way of specifying a stationary point process, namely via *cutting and stacking*. For now, we just note that a point process X with distribution μ is stationary if and only if the fidi distributions of μ coincide with the fidi distributions of $\tilde{S}_t(\mu)$, for all $t \in \mathbb{R}^d$.

Percolation theory is concerned with infinite objects and hence only makes sense on infinite graphs. We require therefore that our percolation models are based on point processes with the property that $X(\mathbb{R}^d) = \infty$. This, however, basically is a consequence of stationarity as we now show.

Proposition 1.1 *Let X be a stationary point process for which*

$$P(X(\mathbb{R}^d) = 0) = 0.$$

Then $P(X(\mathbb{R}^d) = \infty) = 1$.

Proof Suppose that there exists an integer k such that $P(X(\mathbb{R}^d) = k) > 0$. Then there must also exist an integer b such that

$$P\left(X(B_b) > \tfrac{1}{2}k,\, X(\mathbb{R}^d \backslash B_b) < \tfrac{1}{2}k\right) =: \epsilon > 0,$$

where B_b is the set $[-b, b]^d$. Let $r \in Z^d$ be a vector with integer-valued coordinates and let $br = (br_1, \ldots, br_d)$. Consider the events

$$E_r = \left\{X(T_{br}(B_b)) > \tfrac{1}{2}k,\, X(T_{br}(\mathbb{R}^d \backslash B_b)) < \tfrac{1}{2}k\right\}.$$

It follows from the stationarity of X that $P(E_r) = \epsilon$, for all $r \in Z^d$. But the events E_r are disjoint for distinct r, and this is the required contradiction. □