Tutorat de Mathématiques

Tutorat 3: Percolation

On s'intéresse à la description probabiliste du phénomène de percolation sur réseau, ainsi qu'à la diffusion dans un système désordonné.

I Travail préliminaire

1. Lire le texte en annexe, puis présenter succinctement quelques exemples pratiques de percolation.

II Vocabulaire

Considérons un réseau de taille infinie pour lequel les *sites* peuvent être *occupés* avec une probabilité p, ou *inoccupés* avec une probabilité $1-p$. On appelle *amas de masse* s tout ensemble de s premiers voisins occupés.

On appelle *seuil de percolation* la probabilité $p = p_c$ au delà de laquelle il existe un *amas percolant*, c'est à dire un amas rejoignant les extrémités du système.

On définit également la *taille caractéristique des plus grands amas* ξ, la fraction des sites dans l'amas infini ou *paramètre d'ordre* $\mathcal{P}$, et enfin la *masse moyenne des amas finis* $\mathcal{M}$. On note ν, β et γ leurs exposants critiques respectifs au voisinage du seuil de percolation:

$$\xi \sim |p - p_c|^{-\nu} \tag{1}$$
$$\mathcal{P} \sim |p - p_c|^{\beta} \tag{2}$$
$$\mathcal{M} \sim |p - p_c|^{-\gamma} . \tag{3}$$

III Percolation à une dimension

Afin de bien poser les bases et le vocabulaire, on étudie dans un premier temps le cas trivial d'un réseau infini à une dimension (voir Fig. 1).

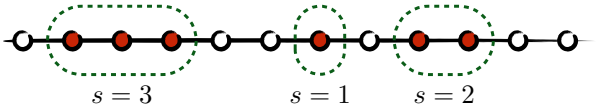


Figure 1: Réseau 1D. Les sites occupés sont signalés en rouge.

2. Montrer que la probabilité $\pi(s)$ qu'un site quelconque appartienne à un amas de masse s s'écrit:

$$\pi(s) = s\,p^s(1-p)^2 . \tag{4}$$

3. En déduire que la densité $n(s)$ des amas de masse s (c'est à dire le nombre d'amas de mass s par site) s'écrit:

$$n(s) = p^s(1-p)^2 . \tag{5}$$

4. Que vaut ici le seuil de percolation p_c ?

5. Afin d'étudier plus précisément le comportement critique au voisinage de p_c, on pose $p = p_c - \varepsilon$ avec $\varepsilon \ll 1$. Montrer que l'on a:

$$n(s) \sim \varepsilon^2 e^{-s\varepsilon} . \tag{6}$$

puis déterminer la valeur de l'exposant critique ν.

6. On définit communément la masse moyenne des amas finis par la relation $\mathcal{M} = \sum_{s=0}^{\infty} s\,\pi(s)$. Montrer que l'on a:

$$\mathcal{M} \sim \varepsilon^2 \sum_{s=0}^{\infty} s^2 e^{-s\varepsilon} , \tag{7}$$

puis déterminer la valeur de l'exposant critique γ. On pourra penser à transformer la somme discrète en une somme continue.

7. Que dire de l'exposant critique β ?

IV Percolation à deux dimensions

On s'intéresse dans un premier temps au cas du réseau carré (voir Fig. 2).

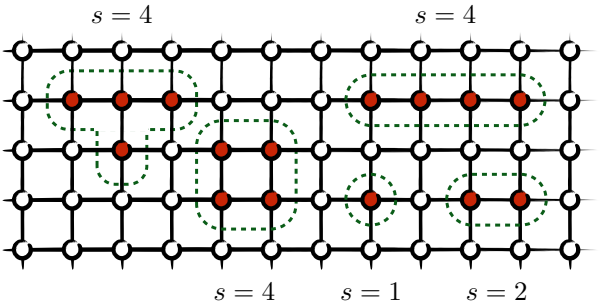


Figure 2: Réseau carré.

8. Montrer que la probabilité $\pi(s)$ qu'un site quelconque appartienne à un amas de masse s s'écrit maintenant:

$$\pi(s) \quad = \quad \sum_t s\, p^s (1-p)^t g_{st} \ , \tag{8}$$

où t est le nombre de sites vides encadrant l'amas de taille s, et g_{st} est facteur de dégénérescence dont on donnera le sens. On pourra traiter les cas $s = 3$ et $s = 4$ en guise d'exemple.

A ce jour, on ne sait pas calculer analytiquement le facteur g_{st}. Nous allons donc regarder un réseau particulier, dit *de Bethe*, dont la topologie nous permettra d'aller plus loin dans l'étude de la percolation à deux dimensions (voir Fig. 3). Le réseau de Bethe est un réseau *arborescent*, c'est à dire qu'il existe un unique chemin reliant deux sites donnés du réseau. Le réseau de Bethe n'est défini que par sa coordinence z, il n'a pas de métrique et donc le calcul de la quantité ξ n'a aucun sens. Enfin, étant donné un site du réseau, on appelle *branche* l'ensemble des sites qui lui sont connectés par l'un de ses z bras.

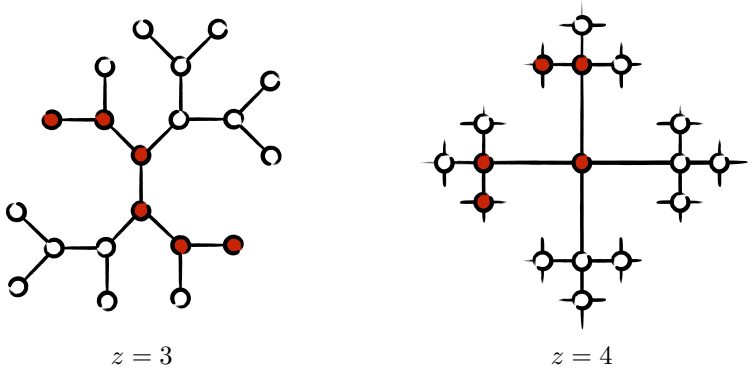


Figure 3: Réseaux de Bethe.

Il est nécessaire ici de poser à nouveau la définition de percolation: on dit d'une branche du réseau qu'elle percole dès lors que son premier site est occupé *et* qu'elle porte un amas infini.

9. Etablir une relation polynomiale sur la probabilité $\mathcal{Q}$ qu'une branche ne percole pas faisant intervenir p et z.

10. Dans le cas $z = 3$, déterminer la fonction $\mathcal{Q}(p)$ puis la tracer pour $p \in [0, 1]$.

11. Dans le cas général, montrer que le seuil de percolation est donné par:

$$p_c \quad = \quad \frac{1}{z-1} \ . \tag{9}$$

Pour ce faire, on pourra développer la relation obtenue à la question **9.** au voisinage de p_c.

12. La paramètre d'ordre pouvant être vu comme la probabilité qu'un site appartienne à l'amas infini, déterminer l'expression de $\mathcal{P}$ en fonction de $\mathcal{Q}$, z et p.

13. Déterminer l'exposant β. On pourra à nouveau utiliser le développement au voisinage de p_c de la relation obtenue à la question **9.**

14. Montrer que la masse moyenne $\mathcal{M}$ de l'amas porté par un site se met sous la forme:

$$\mathcal{M} \quad = \quad p\left(1+\frac{zp}{1-p(z-1)}\right) \ . \tag{10}$$

On pourra dans un premier temps chercher à calculer la masse moyenne de l'amas porté par une branche.

15. Déterminer l'exposant γ.

V Point critique et invariance d'échelle

Au point critique $p = p_c$, les systèmes de percolation sont *invariants d'échelle* ou *fractales*. C'est à dire qu'aucune échelle ne caractérise le système, ou encore que les propriétés du système seront les mêmes quelle que soit la "distance" à laquelle on se place pour l'observer (voir Fig. 4).

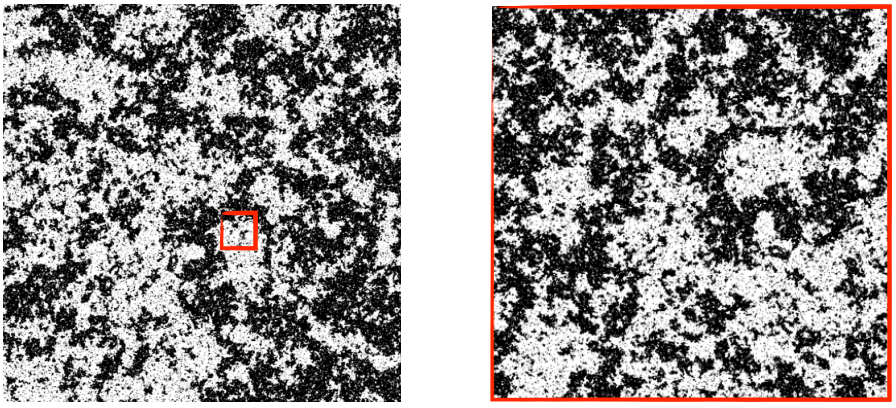


Figure 4: Simulation de percolation au point critique $p = p_c$ à gauche, zoom $\times 10$ à droite. "FK-Ising model (Fortuin-Kasteleyn)" aussi connu comme "random-cluster model".

Définition Une fonction f est dite *invariante d'échelle* dès lors qu'il existe une fonction g telle que:

$$\forall\, x, y \qquad \frac{f(x)}{f(y)} \quad = \quad g\left(\frac{x}{y}\right) \ . \tag{11}$$

16. Démontrer qu'une fonction f est invariante d'échelle *si et seulement si* il s'agit d'une fonction puissance.

Pour la suite de cette partie, on se place dans le cas particulier du réseau de Bethe de coordinence $z = 3$.

17. Montrer que la densité des amas de masse s s'écrit:

$$n(s) \quad = \quad g_s\, p^s (1-p)^{s+2} \ , \tag{12}$$

où g_s est une facteur de dégénérescence dont on donnera le sens.

18. On note $n_c(s) \sim s^{-\tau}$ la densité des amas de masse s au point critique (τ est appelé *exposant de Fisher*). Donner le comportement de g_s en fonction de s, p_c et τ.

19. Montrer qu'au voisinage du point critique et aux grands s on a:

$$n(s) \quad \sim \quad s^{-\tau} e^{-4(p-p_c)^2 s} \ . \tag{13}$$

20. Tracer $n(s)$ au voisinage du point critique et déterminer la masse caractéristique des plus grands amas.

21. Déterminer l'exposant τ. On pourra chercher à recalculer $\mathcal{M}$ de la manière suggérée à la question **6.**

VI Jeux de fourmis

On s'intéresse à la diffusion de fourmis dans un réseau désordonné. A chaque pas de temps t, la fourmi choisit l'un des sites plus proches voisins. Si le site est occupé (autorisé) alors elle y va, si le site est inoccupé (interdit) alors elle reste là ou elle est jusqu'au pas de temps suivant. Ce processus est répété de nombreuses fois, puis moyenné sur un grand nombre de fourmis et sur un grand nombre de réseaux. Les interactions fourmi-fourmi sont négligées.

On cherche à déterminer la distance moyenne $\mathcal{R}$ parcourue par une fourmi en fonction du temps. On note $\mathcal{R} \sim t^k$. Pour $k = 1/2$ on parle de diffusion *normale*, alors que pour $k \neq 1/2$ on parle de diffusion *anormale*.

22. A votre avis, quel type de diffusion a t'on pour $p < p_c$? pour $p \gg p_c$? et pour $p = p_c$?

On note $P_i(t)$ la probabilité conditionnelle que la fourmi se trouve au site i à l'instant t sachant que $P_0(0) = 1$. On note σ_{ij} la probabilité que la fourmis saute du site i au site j en un pas de temps.

23. Etablir la loi de propagation de $P_i(t)$ ou encore *equation maitresse:*

$$P_i(t+1) - P_i(t) \quad = \quad \sum_{j|\langle i,j\rangle} \left[\sigma_{ji} P_j(t) - \sigma_{ij} P_i(t)\right] . \tag{14}$$

24. Que vaut σ_{ij} pour une fourmi ivrogne (ou aveugle) ? et pour une fourmi sobre ? On note z la coordinence du réseau et z_j le nombre de plus proches voisins occupés du site j.

25. Argumenter que dans un amas fini, on finit par atteindre un état stationnaire pour lequel $P_i(t)$ ne dépend plus du temps.

Pour $p < p_c$ proche du seuil de percolation et aux temps longs, on peut montrer que R est indépendant du temps et se comporte comme $R \sim (p_c - p)^{\beta/2-\nu}$. Par ailleurs, pour $p > p_c$ on a aux temps longs $R^2 \sim \mathcal{D}t$ (diffusion normale). Enfin, proche du seuil de percolation, la diffusivité se comporte comme $\mathcal{D} \sim (p - p_c)^{\mu}$.

Afin de déterminer la loi de diffusion au seuil de percolation, on pose $R = t^k f\left[(p - p_c)t^m\right]$.

26. Montrer que $f_{p>p_c}[z] \sim z^{\mu/2}$, et que $f_{p<p_c}[z] \sim (-z)^{-k/m}$.

27. En déduire qu'au seuil de percolation $p = p_c$, la loi de diffusion est donnée par:

$$R \quad \sim \quad t^{(\nu-\beta/2)/(2\nu+\mu-\beta)} . \tag{15}$$

28. En dimension 2, des méthodes analytiques avancées permmettent de déterminer $\beta = 5/36$, $\nu = 4/3$, et des simulations numériques donnent $\mu/\nu \simeq 0.975$. Conclure sur le type de diffusion.

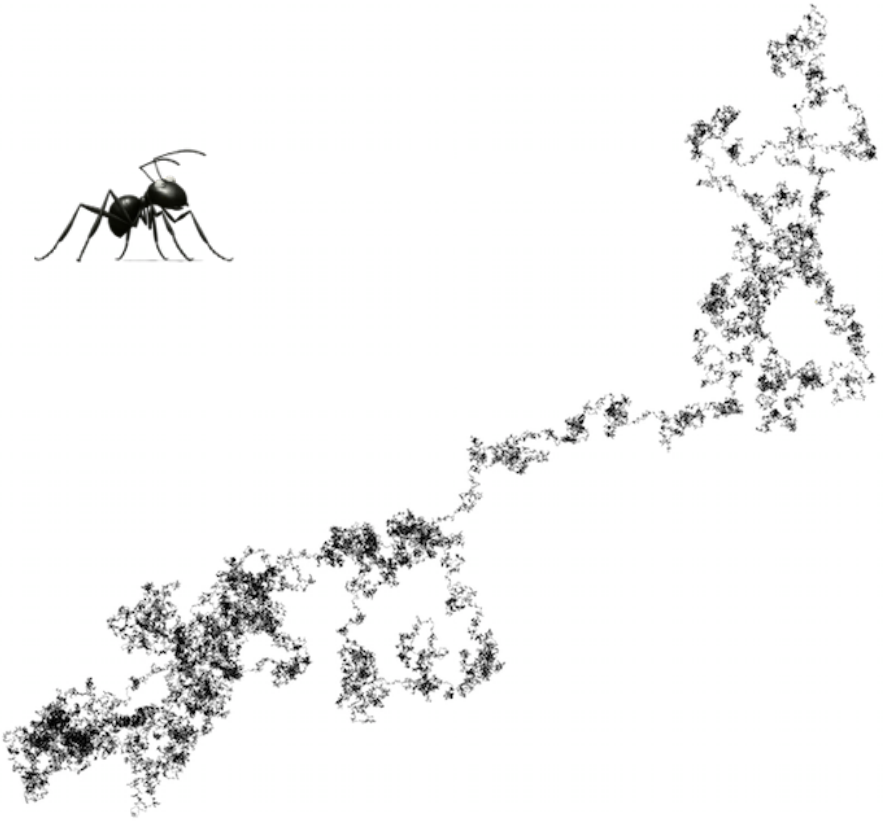


Figure 5: Marche aléatoire sur π. 10^5 pas de temps.

1.1. WHAT IS PERCOLATION?

Imagine a large array of squares as shown in Fig. 1(a). We imagine this array to be so large that any effects from its boundaries are negligible. Physicists call such an array a square lattice, mathematicians denote it by $\mathbb{Z}^2$; common sense identifies it with a big sheet of ruled paper. (You may complain that the square lattice in Fig. 1(a) is not very large, but the publisher did not allow us to fill all remaining pages of this book with these squares, which would have greatly simplified our task of writing the book and yours of reading it.) Now a certain fraction of squares are filled with a big dot in the centre, whereas the other squares are left empty, as in Fig. 1(b). We now define a *cluster* as a group of neighbour squares occupied by these big dots; these clusters are encircled in Fig. 1(c). From this picture we see that squares are called nearest neighbours if they have one side in common but not if they only touch at one corner. Physicists call squares with one common side 'nearest neighbour sites on the square lattice', whereas squares touching at one corner only are 'next nearest neighbours'. All sites within one cluster are thus connected to each other by one unbroken chain of nearest-neighbour links from one occupied square to a neighbour square also occupied by a big dot. The graphical 'cluster' explanation through Fig. 1(c) seems more appropriate for our purposes here than a precise mathematical definition. Percolation theory now deals with the number and properties of these clusters; perhaps the reader will agree with us that there are not many requisites needed to understand what percolation theory is about.

How are the dots distributed among the squares in Fig. 1? One may assume that the dots love to cling together, or that they hate each other and try to move as far away as possible. But the simplest assumption is that they ignore each other, not unlike scientists working in similar fields. Then the occupation of the squares is *random*, that is each square is occupied or empty independent of the occupation status of its neighbours. We call p the probability of a site being occupied by a big dot; that means that if we have N

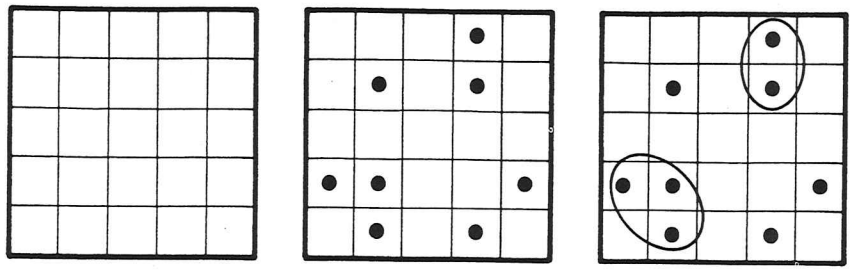


Fig. 1. Definition of percolation and its clusters: (a) shows parts of a square lattice; in (b) some squares are occupied with big dots; in (c) the 'clusters', groups of neighbouring occupied squares, are encircled except when the 'cluster' consists of a single square.

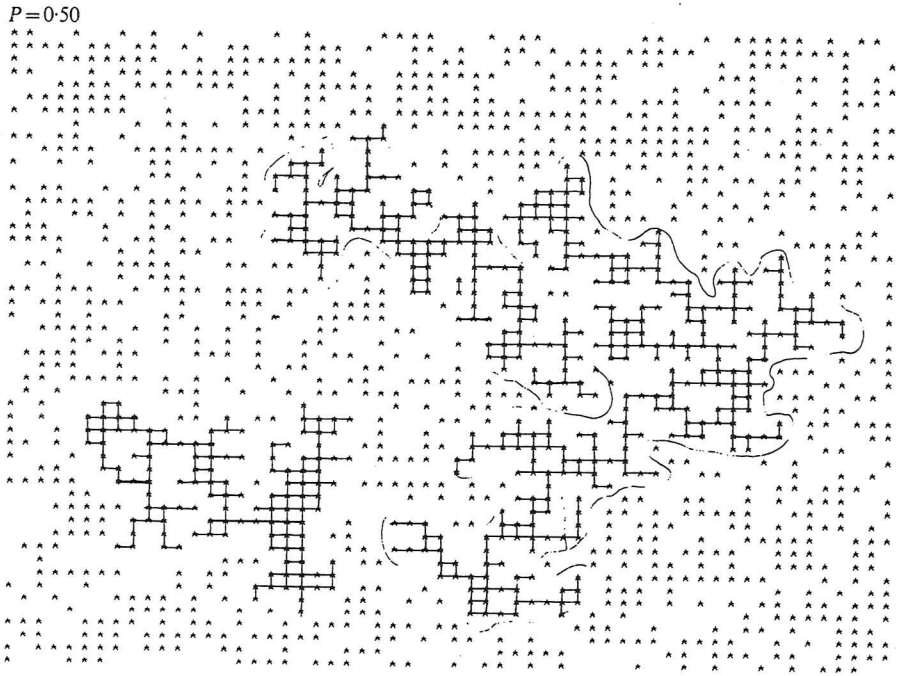
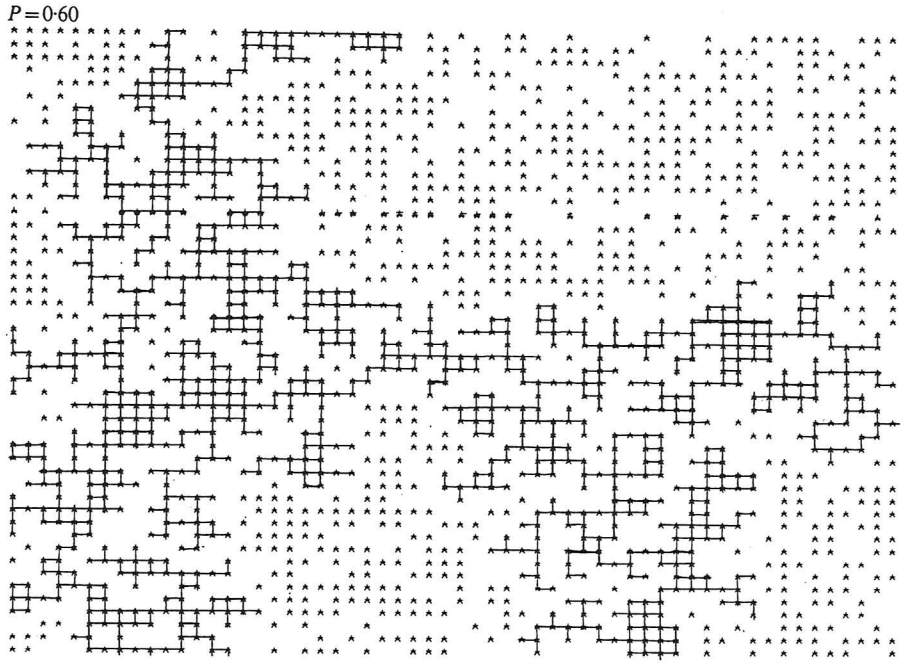
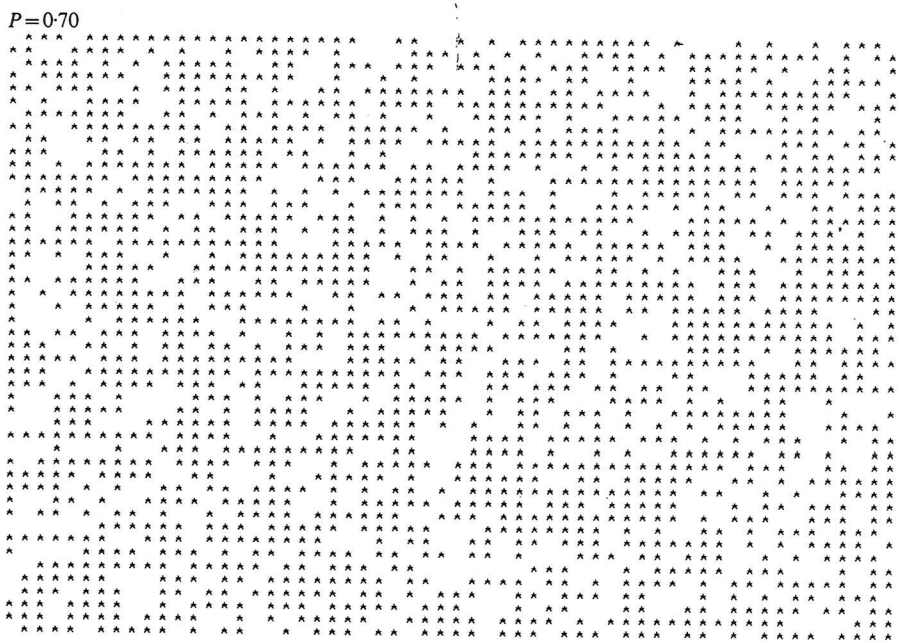


Fig. 2. Example for percolation on a 60×50 square lattice, for various p as indicated. Occupied squares are shown as *, empty squares are ignored. Near the threshold concentration 0·5928 the largest cluster is marked.





squares, and N is a very large number, then pN of these squares are occupied, and the remaining $(1-p)N$ of these squares are empty. This case of random percolation is what we concentrate on here:

Each site of a very large lattice is occupied randomly with probability p, independent of its neighbours. Percolation theory deals with the clusters thus formed, in other words with the groups of neighbouring occupied sites.

Of course, the reader may replace 'occupied by a big dot' with 'black' and 'empty' with 'white' (or 'red', if he likes politics); or he may use any other suitable pair of words denoting two mutually exclusive states of the site.

Figure 2 shows a computer-generated sample of a 60×50 square lattice, with probability p increasing from 10% to 90%. We see that for $p \geqslant 0{\cdot}6$ one cluster extends from top to bottom and from left to right of the sample; one says that this cluster percolates through the system rather like water percolates through a coffee machine. A large part of this book deals with the peculiar phenomena of percolation near that concentration p_c where for the first time a percolating cluster is formed. These aspects are called *critical phenomena*, and the theory attempting to describe them is the *scaling theory*.

Historically, percolation theory goes back to Flory and to Stockmayer who during World War II used it to describe how small branching molecules form larger and larger macromolecules if more and more chemical bonds are formed between the original molecules. This polymerization process may lead to *gelation*, that is to the formation of a network of chemical bonds spanning the whole system. Thus the original small molecules correspond to our squares, the macromolecules to our clusters, and the network to our percolating cluster. You may be an experienced researcher in percolation without having been aware of it, for the boiling of an egg, which is first liquid and then becomes more solid-like ('gel') upon heating is an example. Flory and Stockmayer developed a theory which today one calls percolation theory on the Bethe lattice (or Cayley tree) and which will be explained later. But until recently it was controversial whether critical phenomena for gelation are described correctly by percolation theory and its assumption that chemical bonds are formed randomly (de Gennes, 1976; Kolb and Axelos, 1990).

Usually, the start of percolation theory is associated with a 1957 publication of Broadbent and Hammersley which introduced the name and dealt with it more mathematically, using the geometrical and probabilistic concepts explained above. Hammersley, in his personal history of percolation in *Percolation Structures and Processes*, mentions that the new computers which became available to scientists at that time were one of the reasons for developing percolation theory as a problem where the computers could be useful. We will see later that even today computers play a crucial role for percolation, with lattices containing thousands of millions of sites being simulated and analysed.

The percolation theory as described here, with its particular emphasis on critical phenomena, was developed since the 1970s; one may regard a note by Essam and Gwilym in 1971 as one of the starting points of the later avalanche of publications. Instead of going through the details now we describe three simple 'games' which can be easily simulated on a computer and which may serve as an introduction to a reader preferring to learn percolation by a 'hands-on' approach. These example are somewhat unusual, and the reader may skip them and proceed with Chapter 2.

1.2. FOREST FIRES

This section introduces a simple model for forest fires. Its aim is not so much to help fighting fires but to help to understand the idea of a percolation threshold, the concept of a sharp transition with diverging times, and computer simulation.

French scientists in Marseilles and elsewhere are interested, for obvious reasons, in understanding and controlling forest fires. They told us of the following percolation problem which can easily be simulated on a computer. How long does a forest fire take to either penetrate the forest or to be extinguished?

As is well known, a diligent student should make hundreds of independent experiments to reduce statistical errors before reporting the results in his thesis. If for every thesis, a hundred fires were initiated in the forests surrounding the university, society's respect for research might be diminished. It is much more practical to simulate numerous such fires on a computer. For this purpose we approximate the forest by a square lattice. Each square in Fig. 1 is either occupied by a tree, in which case we call that site 'green', or it is empty, in which case we call it 'white'. The probability for a green square is p, that for a white square is $(1-p)$. For $p=1$ all squares would correspond to trees, which would be appropriate to a garden of apple trees but not for a natural forest. The fact that $p<1$ allows for holes (white squares) which cause disorder in the forest. This distribution of white and green sites (squares) is our initial state.

Now let some trees burn and call those squares which correspond to burning trees 'red' sites. The simplest choice is to light all the trees in the first row of the lattice, whereas the remaining trees, in lines $2, 3, \ldots, L$ of the $L \times L$ lattice, remain green. Does this fire on one side of the forest penetrate through the whole forest down to line L of our array?

For this purpose we have to clarify how a tree can ignite the other trees. To simplify the computer simulation we go through our lattice regularly, first scanning the first line of trees from left to right and checking which neighbours they ignite, then scanning the second line in the same way, and so on until we reach the last line of trees. During the whole simulation, a green tree is ignited and becomes red if it neighbours another red tree which at that time is still burning. Thus a just-ignited tree ignites its right and bottom neighbour

within the same sweep through the lattice, its top and left neighbour tree at the next sweep. Reaching the end, we start again with the tree at the extreme left in the first line. Each sweep through the whole lattice (experts call that one Monte Carlo step per site) constitutes one time unit in our simulation. We assume that the fire can spread only to green nearest neighbour trees, not to trees which are farther away. Furthermore, a tree which has burnt during one time unit is regarded as burnt out ('black') and no longer ignites any other tree. We regard the forest fire as terminated if it either has reached the last line or if no burning trees are left. (In the first case, the fire would ignite the next line of trees if a larger lattice had been stored in the computer; in the second case, only black trees and green trees adjoining white places are left over, the black trees constituting formerly burning trees which have burnt out, the green trees never having been touched by the fire since they were separated safely from the burning trees). The *lifetime* of the forest fire is defined as the number of sweeps through the lattice until termination is reached, averaged over many distributions of trees among the sites of the same lattice at the same probability p.

Figure 3 shows this lifetime of forest fires as a function of the probability p that a square is occupied with a tree. These simple computer simulations indicate that there is a sharp transition, for the above case near $p = 0{\cdot}6$, where the lifetime seems to approach infinity. Of course, in the simulation of finite lattices the reader cannot expect truly infinite times; but one can simulate the forest fires at the same 'critical' value of p near $0{\cdot}5928$ for different lattice sizes and show that the lifetime increases with increasing size of the forest.

Why is there a special value of p, which we call the percolation threshold p_c, where the lifetime seems to diverge? For p near unity, each row can immediately ignite the trees in the next row, and thus after one sweep through the lattice the fire may already have reached the last row. For p near zero, most burning trees have no neighbours at all, and the fire stops there after the tree has burnt out; thus after a few sweeps nothing burns anymore. If we increase p from small values to large values, then at some critical value $p = p_c$ a path of neighbouring trees appears which connects the top row with the bottom row for the first time, that is we see a percolating cluster. The shortest path which, for p slightly above p_c, this percolating network creates to connect top and bottom is called the minimal, or chemical path. It will in general be very different from a straight line. Fig. 4 shows a typical path. (See also Section 5.2.)

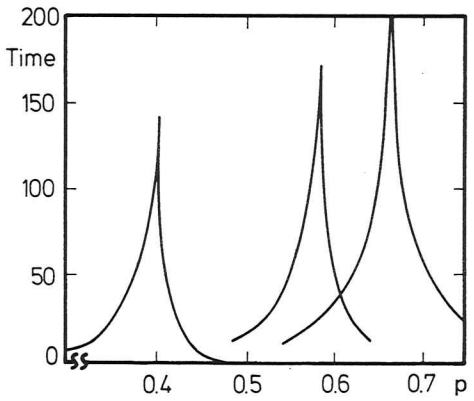


Fig. 3. Average termination time for forest fires, as simulated on a square lattice. The centre curve corresponds to the simplest case described in the text. The lefthand curve gives data if the fire can spread to both nearest and next nearest neighbours. For the righthand curve two burning trees are needed to ignite a nearest or next-nearest neighbour.

Because of the simplified way in which we construct our model, the fire spreads preferentially from top to bottom, or left to right, and needs a much longer time to move backwards from right to left or from bottom to top. For four consecutive forward steps, say top to bottom, it needs just one time unit, whereas four backward steps require four time units, as the reader can easily check on this figure by going through the above algorithm. Thus now the fire needs a long time to penetrate the forest. If p is diminished to a value slightly below p_c, then some trees, for example the one marked by an X in Fig. 4, may be missing. The fire then needs a long time to find out that it cannot penetrate the forest, and thus only after many sweeps through the lattice will the fire be extinguished. Therefore the lifetime will become very large if p approaches p_c from below or above.

We also show in Fig. 3 the results for two modifications of the above model. In one case we allow the fire to spread not only to the nearest neighbour trees (squares which have one side in common) but also to next-nearest neighbours (squares which have only one corner in common). Then the critical

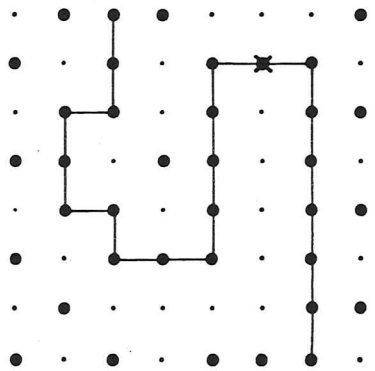


Fig. 4. Example of the shortest path connecting the top line of a small square lattice with the bottom line, for p slightly above p_c. The straight sections of this line connect the centres of occupied squares, the X marks the site which, if missing due to a small reduction of p, would disconnect top and bottom lines but would still give a long termination time for the forest fire simulation.

point is shifted to about 0·4; experts have shown that it is at one minus the above critical value, i.e. $1 - 0{\cdot}5928 = 0{\cdot}4072$. But even without much thinking and computing one can understand that now the fire can spread more easily since it can jump over longer distances; therefore the percolation threshold is lowered.

The other modification goes in the opposite direction: We assume that the weather is more like in Nova Scotia (Canada) than in Marseilles (France). Since it is quite cold, a tree needs two burning neighbours, instead of only one, before it can ignite. Now it is more difficult for the fire to percolate through the forest, and the percolation threshold is shifted upwards, as the simulations in Figure 3 show.

The reader may complain that the above algorithm gives the fire a preference to spread to the right and the bottom and may dislike these similarities with political or economic trends, respectively. But for forest fires, such trends can be justified as representing a wind blowing in one 'diagonal' direction. In reality this preference is introduced to save computer time.

For readers interested in the physics of phase transitions it should be mentioned that the percolation threshold at $p = p_c$ gives the position of a phase transition (for experts only: without 'broken symmetry'). At a phase transition, a system changes its behaviour qualitatively for one particular value of a continuously varying parameter. In the percolation case, if p increases smoothly from zero to unity, then we have no percolating cluster for $p < p_c$ and (at least) one percolating cluster for $p > p_c$. Thus at $p = p_c$, and only there, something peculiar happens: for the first time a path of neighbouring green trees connects top and bottom. Also the divergence of characteristic times (in our case the fire lifetime) at the critical point has analogies in other phase transitions where it is called '*critical slowing down*'. For example, for a temperature only slightly below the liquid–gas critical temperature, the fluid is quite unsure whether it wants to be liquid or vapour, and thus takes a lot of time to make its choice; this time can be measured by light scattering. Similarly, relaxation times near magnetic Curie points are very large.

1.3. OIL FIELDS AND FRACTALS

Percolation can be used as an idealized simple model for the distribution of oil or gas inside porous rocks in oil reservoirs. In Fig. 1, imagine that the unoccupied (white) squares represent regions filled with hard rock, while the occupied squares represent pores that are filled with oil or gas. The average concentration of oil in the rock is represented by the occupation probability p. (In the oil terminology, p is called 'porosity'.) In real reservoirs, the mechanisms that created the oil deposits imply some *correlations* between occupied pores, owing to the way the rock was originally cracked or the way the different deposits were put in place. The simple percolation model ignores these correlations, and assumes that each basic square (or cube, if this is repeated in three dimensions) is occupied or empty independently of its neighbours. However, the qualitative features described below also hold in the more realistic models.

It is obvious from Fig. 2, that when p is smaller than p_c, the oil is found only in finite connected clusters. Therefore if we place a well at a random site, it will most probably hit a small cluster, produce a finite small amount of oil and be a very bad investment. To produce a large amount of oil, we need a reservoir which has $p > p_c$, and we need to have the well at a site that belongs to the largest cluster.

The oil people are very interested in predicting how much oil they would produce from a well. To help in these predictions, they take out rock samples from the well. These come in long rock logs, with a typical diameter of order 5–10 cm. One can then measure the porosity (percentage of pores) in a piece of linear size 5 cm, and try to *extrapolate* to the reservoir scale, which could be many kilometres. Is such extrapolation legitimate?

To address this question, let us identify our 'well' as one site sitting in the square example of Fig. 2 and belonging to the largest cluster for $p > p_c$. Let us next put a frame of size $L \times L$ around this point, and count how many points within this frame belong to the same cluster, $M(L)$. The reader can easily try this exercise with frame sizes $L = 3, 5, 7, 9$, etc. Looking at the last example in Fig. 2, it is clear that $M(L)$ practically grows linearly with the area of the frame, L^2, and we can define the average density of points connected to our well as $P = M(L)/L^2$. P is then independent of L, and is monotonically decreasing as p decreases. However, the situation is very different for p very close to p_c, e.g. $p = 0{\cdot}6$. In that case, the largest cluster is rather *ramified*, and it has many 'holes' in it. Those holes contain other clusters, which may be quite large, but whose oil is not reachable through our well. Looking at the picture for $p = 0{\cdot}6$ in Fig. 2, one sees 'holes' *on many length scales*. As we shall see later, the occurrence of phenomena on all length scales is very basic for many of the interesting phenomena, which occur near p_c.

If one measures $M(L)$ as function of L *at* p_c, the result is no longer linear in the area L^2. In fact, if one plots log $M(L)$ versus log L for bigger lattices (see also Fig. 15), one finds a straight line with slope 1·9, implying that at p_c one has

$$M(L) \propto L^{1{\cdot}9}$$

(We use the symbol $\propto$ to indicate proportionality. In many cases this proportionality is meant to be accurate only in the *asymptotic* limit, here of large L.) The exponent 1·9 is called a 'fractal dimensionality', or fractal dimension.

You may have noticed with sadness that a small bottle of scotch, half as high as the customary whisky bottle, does not contain half as much of the precious fluid but only one eighth; not only the height is reduced by a factor 2 but also the width and the depth, with the volume being the product of these three lengths. In other words, a bottle is a three-dimensional object. For two

dimensions, a piece of paper which has a length and a width both twice as large as those of another piece weighs four times as much. Only a one-dimensional object, like a long wire, is simple. A wire half the length of another weighs half as much as the longer wire. In all these cases, the mass M scales with the linear size L as $M \propto L^d$, and d is the usual *Euclidean dimension*. Benoit Mandelbrot introduced 'fractal geometry' as a unifying description of natural phenomena which are not uniform but still obey simple power laws of the form

$$M \propto L^D$$

with non-integer dimensions D. For three-dimensional percolation clusters at p_c one finds $D \simeq 2{\cdot}5$. We shall see more examples of *fractals* below.

The fact that $M(L)$ grows as $L^{1{\cdot}9}$ implies that the average density $M(L)/L^2$ is not constant, but rather decays as $L^{-0{\cdot}1}$. Therefore, the average density of the extractable oil in a field with porosity near p_c, of size 100 km, is smaller by a factor of about $(10^6)^{-0{\cdot}1} \simeq 0{\cdot}25$ than that of a sample of size 10 cm. The remaining 75 per cent is not directly connected to the drill hole. Such a factor is crucial if we are to base the economy of oil production on it! The corresponding factor in three dimensions is $(10^6)^{-0{\cdot}5} = 10^{-3}$!

In fact, the situation is not so bad, since the density does become uniform for large L above p_c. As we shall see, there exists a typical length $\xi(p)$, called the *correlation length*, such that $M(L) \propto L^{1{\cdot}9}$ for $L < \xi$, and $M(L) \propto L^2$ for $L > \xi$. ξ is a measure of the largest hole in the largest cluster, and it decreases as we increase p above p_c. However, the oil people should use a sample larger than ξ in order to estimate the correct amount of oil they can get. A more quantitative discussion of this problem will be given below, in Section 3.4.

The problem of extracting oil from the rock involves not only estimating the amount of such oil, but also discussing the *flow* of the fluid inside the porous medium. This brings up many questions concerning *dynamics* on the percolation clusters, that we shall discuss below. The simplest example, concerning *diffusion*, is briefly introduced in the next section.

The reader should be warned, however, that both these remarks on oil flow, as well as the earlier ones on forest fires, are meant as illustrations, not as proven engineering applications.

1.4. DIFFUSION IN DISORDERED MEDIA

Hydrogen atoms are known to diffuse through many solids, an effect which might become important for energy storage. If the solid is not a regular lattice, this diffusion takes place in a disordered, not an ordered medium. A particularly simple disordered medium is our percolation lattice, where only a fraction p of all sites (squares) is occupied, the rest are empty. Let us assume the hydrogen atom can move only from one occupied site of the lattice to a nearest neighbour which is also occupied. Then the motion is restricted to the cluster of percolation theory to which the atom belongs initially. It can never jump to another cluster since then it would have to move at least once over a distance larger than that between nearest neighbours. This problem was called the '*ant in the labyrinth*' by de Gennes in 1976. At the beginning of the 1980s this problem became very fashionable, particularly at the percolation threshold $p = p_c$.

Another useful application of this would concern the diffusion of test particles through the oil in the porous rock, mentioned above. Such diffusion is sometimes used to study the properties of the pore structure.

Let us not care whether hydrogen atoms move through solids, an ant tries to escape a labyrinth, or the reader desperately searches for a way through this book. We simply have a point, called an ant, which sits on an occupied square of our square lattice and which at every time unit makes one attempt to move. This attempt consists in randomly selecting one of its four neighbour squares. If that square is occupied, it moves to that square; if instead it is empty, the ant stays at its old place. In both cases the time t is increased by one unit after the attempt. After a certain time t, one calculates the squared distance between the starting point and the end point. One repeats the simulation by giving the ant a different occupied square as a starting point; finally, one averages the squared distance obtained in this way over many ant movements on many lattices at the same p and same lattice size. How does R, the square root of this averaged squared distance (also called the *root mean square* or rms displacement) depend on time t?

For $p = 1$ one has diffusion on a regular lattice without disorder, and elementary statistical considerations give $R^2 = t$ exactly, if our squares have a length equal to unity. (Proof: For each such random walk, the end-to-end vector R is the vector sum of t displacement vectors d_i, $i = 1, 2, \ldots, t$. When we calculate the square of that sum and then its average, we have to calculate the averages of the scalar products $d_i d_j$. For $i = j$, this scalar product is simply the square of the nearest neighbour distance, which is unity. For i and j different, the scalar product can be $+1$ or -1 with equal probability since we assumed that the motion is completely random. Moreover, in half of the cases the scalar product is zero since d_i and d_j are perpendicular to each other. Thus on average this product cancels out except for $i = j$ where it gives unity. Therefore the squared sum equals t. This proof is not necessary to understand the remainder of the book since we will mainly deal with problems which are not exactly solved.)

Figure 5 shows the results of simple computer simulations on the square lattice. On this double-logarithmic plot one sees the power law $R = \text{const} \times t^k$ more easily than on a normal diagram. It seems to describe the relation between distance R and time t for sufficiently large t. Since $\log R = \log(\text{const}) + k \log(t)$, power laws show up as straight lines in such log-log plots, with the slope giving the exponent k of the power law. We see that for a concentration p far above $p_c = 0{\cdot}59$, k is near $1/2$ for large times, whereas for p far below p_c the distance R approaches a constant for large

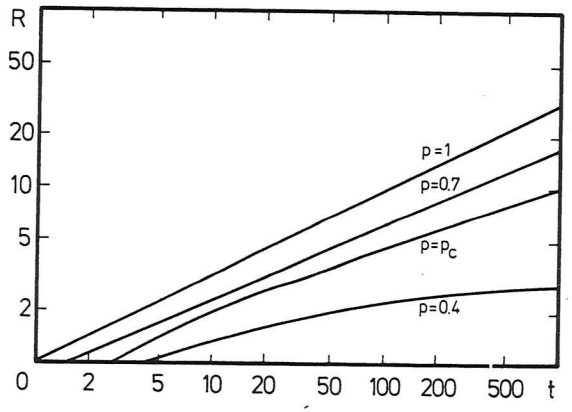


Fig. 5. Example of the distance R travelled by an ant in a labyrinth, as simulated on a square lattice for $p < p_c$, $p = p_c = 0{\cdot}5928$, and $p > p_c$. Note the double-logarithmic scales.

times, that is $k = 0$. Right at p_c, k takes on a value in between these two extremes and is roughly equal to $1/3$. This effect, that the exponent is neither that for normal diffusion ($k = 1/2$) nor that for a constant distance ($k = 0$) was called anomalous diffusion by Gefen, Aharony and Alexander. Again it has an analogue in critical points near thermal phase transitions. For example spin diffusion in ferromagnets at $T = T_c$ or mass diffusion at liquid–gas critical points no longer follows the normal diffusion laws but is often described by an anomalous diffusion exponent k or $1/z$.

It is easy to understand why the ant moves so differently for p above and below the percolation threshold p_c. For $p < p_c$ there are only finite clusters present, and the ant sits on one of them. Thus it moves only within that cluster (if this cluster happens to be an isolated occupied square, the ant cannot move at all.) Therefore its motion is restricted over finite distances and R approaches a value connected to the cluster radius if t is very large. For $p > p_c$, on the other hand, the ant can move to infinity if it starts on the percolating network. There are certain holes in this network; but for distances larger than the typical hole size, the ant feels only an average over the small holes, just as the tyres of your Rolls-Royce average over the small pores of the asphalt over which your chauffeur is driving you. Thus the disorder acts as a friction which slows down the diffusion process but does not prevent it $k = 1/2$ for long times. Only at the border case $p = p_c$, does the ant not know which of the two power laws it should follow.

Considering t as the number of steps the ant performs, and R as the linear size of the region visited by the ant, the relation $t \propto R^{1/k}$ can also be interpreted as stating that the number of steps in a region of linear size R is *fractal* with a fractal dimension equal to $1/k$. For regular lattices, and in the homogeneous regime describing the largest percolation cluster on large length scales above p_c, this exponent $1/k$ is equal to 2. *At* p_c, the ant is restricted to move on clusters which are themselves fractal. It is thus forced to move back and forth within a small piece of the cluster, until it finds its way out. This takes a long time and therefore the number of steps within a restricted area is large, and the fractal dimension of the walk, $1/k$, is larger than 2.

At intermediate concentrations, like $p = 0{\cdot}7$, the ant feels some fractal structure on short distances ($R < \xi$). On those distances, the slope k is close to its value at p_c, i.e. $k \simeq 1/3$. Only when $R \gg \xi$ does the slope approach the uniform value $k = 1/2$. As seen from Fig. 5, the curve for $p = 0{\cdot}7$ has not yet become completely parallel to that of $p = 1$, even at $t = 500$. Better quality data, on larger samples, are needed to confirm the details of the crossover from anomalous to normal diffusion. The low-quality data of Fig. 5 are mainly meant to exhibit results that a student can readily produce on a personal computer. We shall present a more quantitative discussion of diffusion in Chapter 6.