Exercises

12.12 **Percolation and universality.**[1] [2] (Complexity) ④

Cluster size distribution: power laws at p_c. A system at its percolation threshold p_c is self-similar. When looked at on a longer length scale (say, with a ruler with notches spaced $1+\epsilon$ farther apart, for infinitesimal ϵ), the statistical behavior of the large percolation clusters should be unchanged, if we simultaneously rescale various measured properties according to certain rules. Let x be the length and S be the size (number of nodes) in a percolation cluster, and let $n(S)$ be the probability that a given cluster will be of size S at p_c.[3] The cluster measured with the new ruler will have a length $x' = x/(1-\epsilon)$, a size $S' = S/(1+c\epsilon)$, and will occur with probability $n' = (1+a\epsilon)\,n$.

(a) *In precise analogy to our analysis of the avalanche size distribution (eqns 12.3–12.6), show that the probability is a power law, $n(S) \propto S^{-\tau}$. What is τ, in terms of a and c?*

In two dimensions, there are exact results known for many properties of percolation. In particular, it is known that[4] $\tau = 187/91$. You can test this numerically, either with the code you developed for Exercise 2.13, or by using the software at our web site [129].

(b) *Calculate the cluster size distribution $n(S)$, both for bond percolation on the square lattice and for site percolation on the triangular lattice, for a large system size (perhaps $L \times L$ with $L = 400$) at $p = p_c$.*[5] *At some moderate size S you will begin occasionally to not have any avalanches; plot $\log(n(S))$ versus $\log(S)$ for both bond and site percolation, together with the power law $n(S) \propto S^{-187/91}$ predicted by the exact result. To make better use of the data, one should bin the avalanches into larger groups, especially for larger sizes where the data is sparse. It is a bit tricky to do this nicely, and you can get software to do this at our web site [129]. Do the plots again, now with all the data included, using bins that start at size ranges $1 \le S < 2$ and grow by a factor of 1.2 for each bin.* You should see clear evidence that the distribution of clusters does look like a power law (a straight line on your log–log plot), and fairly convincing evidence that the power law is converging to the exact result at large S and large system sizes.

The size of the infinite cluster: power laws near p_c. Much of the physics of percolation above p_c revolves around the connected piece left after the small clusters fall out, often called the *percolation cluster*. For $p > p_c$ this largest cluster occupies a fraction of the whole system, often called $P(p)$.[6] The fraction of nodes in this largest cluster for $p > p_c$ is closely analogous to the $T < T_c$ magnetization $M(T)$ in magnets (Fig. 12.6(b)) and the density difference $\rho_l(T) - \rho_g(T)$ near the liquid–gas critical point (Fig. 12.6(a)). In particular, the value $P(p)$ goes to zero continuously as $p \to p_c$.

Systems that are not at p_c are not self-similar. However, there is a scaling relation between systems at differing values of $p - p_c$: a system coarsened by a factor $1+\epsilon$ will be similar to one farther from p_c by a factor $1+\epsilon/\nu$, except that the percolation cluster fraction P must be rescaled upward by

[1] From *Statistical Mechanics: Entropy, Order Parameters, and Complexity* by James P. Sethna, copyright Oxford University Press, 2007, page 293. A pdf of the text is available at pages.physics.cornell.edu/sethna/StatMech/ (select the picture of the text). Hyperlinks from this exercise into the text will work if the latter PDF is downloaded into the same directory/folder as this PDF.

[2] This exercise and the associated software were developed in collaboration with Christopher Myers.

[3] Hence the probability that a given node is in a cluster of size S is proportional to $Sn(S)$.

[4] A non-obvious result!

[5] Conveniently, the critical probability $p_c = \frac{1}{2}$ for both these systems, see Exercise 2.13, part(c). This enormously simplifies the scaling analysis, since we do not need to estimate p_c as well as the critical exponents.

[6] For $p < p_c$, there will still be a largest cluster, but it will not grow much bigger as the system size grows and the fraction $P(p) \to 0$ for $p < p_c$ as the system length $L \to \infty$.

$1+\beta\epsilon/\nu$.[7] This last rescaling reflects the fact that the percolation cluster becomes more dense as you coarse-grain, filling in or blurring away the smaller holes. You may check, just as for the magnetization (eqn 12.7), that

$$P(p) \sim (p_c - p)^{\beta}. \tag{12.60}$$

In two dimensions, $\beta = 5/36$ and $\nu = 4/3$.

(c) *Calculate the fraction of nodes $P(p)$ in the largest cluster, for both bond and site percolation, at a series of points $p = p_c + 2^{-n}$ for as large a percolation lattice as is convenient, and a good range of n. (Once you get your method debugged, $n = 10$ on an $L \times L$ lattice with $L = 200$ should be numerically feasible.) Do a log–log plot of $P(p)$ versus $p - p_c$, and compare along with the theory prediction, eqn 12.60 with $\beta = 5/36$.*

You should find that the numerics in part (c) are not compelling, even for rather large system sizes. The two curves look a bit like power laws, but the slopes β_{eff} on the log–log plot do not agree with one another or with the theory. Worse, as you get close to p_c the curves, although noisy, definitely are not going to zero. This is natural; there will always be a largest cluster, and it is only as the system size $L \to \infty$ that the largest cluster can vanish as a fraction of the system size.

Finite-size scaling (advanced). We can extract better values for β from small simulations by explicitly including the length L into our analysis. Let $P(p, L)$ be the mean fraction of nodes[8] in the largest cluster for a system of size L.

(d) *On a single graph, plot $P(p, L)$ versus p for bond percolation $L = 5$, 10, 20, 50, and 100, focusing on the region around $p = p_c$ where they differ from one another. (At $L = 10$ you will want p to range from 0.25 to 0.75; for $L = 50$ the range should be from 0.45 to 0.55 or so.) Five or ten points will be fine. You will discover that the sample-to-sample variations are large (another finite-size effect), so average each curve over perhaps ten or twenty realizations.*

Each curve $P(p, L)$ is rounded near p_c, as the characteristic cluster lengths reach the system box length L. Thus this rounding is itself a symptom of the universal long-distance behavior, and we can study the dependence of the rounding on L to extract better values of the critical exponent β. We will do this using a *scaling collapse*, rescaling the horizontal and vertical axes so as to make all the curves fall onto a single scaling function.

First, we must derive the scaling function for $P(p, L)$. We know that

$$\begin{aligned} L' &= L/(1+\epsilon), \\ (p_c - p)' &= (1+\epsilon/\nu)(p_c - p), \end{aligned} \tag{1}$$

since the system box length L rescales like any other length. It is convenient to change variables from p to $X = (p_c - p)L^{1/\nu}$; let $P(p, L) = \bar{P}(L, (p_c - p)L^{1/\nu})$.

(e) *Show that X is unchanged under coarse-graining (eqn 12.61). (You can either show $X' = X$ up to terms of order ϵ^2, or you can show $\mathrm{d}X/\mathrm{d}\epsilon = 0$.)*

The combination $X = (p_c - p)L^{1/\nu}$ is another *scaling variable*. The combination $\xi = |p - p_c|^{-\nu}$ is the way in which lengths diverge at the critical point, and is called the *correlation length*. Two systems of different lengths and different values of p should be similar if the lengths are the same when measured in units of ξ. L in units of ξ is $L/\xi = X^{\nu}$, so different systems with the same value of the scaling variable X are statistically similar. We can turn this verbal assertion into a mathematical scaling form by studying how $\bar{P}(L, X)$ coarse-grains.

(f) *Using eqns 12.61 and the fact that P rescales upward by $(1 + \beta\epsilon/\nu)$ under coarse-graining, write the similarity relationship for $\bar{P}$ corresponding to eqn 12.11 for $\bar{D}(S, R)$. Following our derivation of the scaling form for the avalanche size distribution (through eqn 12.14), show that $\bar{P}(L, X) = L^{-\beta/\nu}\mathcal{P}(X)$ for some function $\mathcal{P}(X)$, and hence*

$$P(p, L) \propto L^{-\beta/\nu}\mathcal{P}((p - p_c)L^{1/\nu}). \tag{12.62}$$

Presuming that $\mathcal{P}(X)$ goes to a finite value as $X \to 0$, derive the power law giving the percolation cluster size $L^2 P(p_c, L)$ as a function of L. Derive the power-law variation of $\mathcal{P}(X)$ as $X \to \infty$ using the fact that $P(p, \infty) \propto (p - p_c)^{\beta}$.

Now, we can use eqn 12.62 to deduce how to rescale our data. We can find the finite-sized scaling

[7] We again assure the reader that these particular combinations of Greek letters are just chosen to give the conventional names for the critical exponents.

[8] You can take a microcanonical-style ensemble over all systems with exactly $L^2 p$ sites or $2L^2 p$ bonds, but it is simpler just to do an ensemble average over random number seeds.

function $\mathcal{P}$ by plotting $L^{\beta/\nu}P(p,L)$ versus $X = (p - p_c)L^{1/\nu}$, again with $\nu = 4/3$ and $\beta = 5/36$.

(g) *Plot $L^{\beta/\nu}P(p,L)$ versus X for $X \in [-0.8, +0.8]$, plotting perhaps five points for each curve, for both site percolation and bond percolation. Use system sizes $L = 5$, 10, 20, and 50. Average over many clusters for the smaller sizes (perhaps 400 for $L = 5$), and over at least ten even for the largest.*

Your curves should collapse onto two scaling curves, one for bond percolation and one for site percolation.[9] Notice here that the finite-sized scaling curves collapse well for small L, while we would need to go to much larger L to see good power laws in $P(p)$ directly (part (c)). Notice also that both site percolation and bond percolation collapse for the same value of β, even though the rough power laws from part (c) seemed to differ. In an experiment (or a theory for which exact results were not available), one can use these scaling collapses to estimate p_c, β, and ν.

[9]These two curves should also have collapsed onto one another, given a suitable rescaling of the horizontal and vertical axes, had we done the triangular lattice in a square box instead of a rectangular box (which we got from shearing an $L \times L$ lattice). The finite-size scaling function will in general depend on the boundary condition, and in particular on the shape of the box.