Continuous phase transitions

12

Continuous phase transitions are fascinating. As we raise the temperature of a magnet, the magnetization will vanish continuously at a critical temperature T_c. At T_c we observe large fluctuations in the magnetization (Fig. 12.1); instead of picking one of the up-spin, down-spin, or zero-magnetization states, this model magnet at T_c is a kind of fractal[1] blend of all three. This fascinating behavior is not confined to equilibrium thermal phase transitions. Figure 12.2 shows the *percolation transition*. An early paper which started the widespread study of this topic [82] described punching holes at random places in a conducting sheet of paper and measuring the conductance. Their measurement fell to a very small value as the number of holes approached the critical concentration, because the conducting paths were few and tortuous just before the sheet fell apart. Thus this model too shows a continuous transition: a qualitative change in behavior at a point where the properties are singular but continuous.

Many physical systems involve events of a wide range of sizes, the largest of which are often catastrophic. Figure 12.3(a) shows the energy released in earthquakes versus time during 1995. The Earth's crust responds to the slow motion of the tectonic plates in continental drift through a series of sharp, impulsive earthquakes. The same kind of crackling noise arises in many other systems, from crumpled paper [61] to Rice Krispies™ [72], to magnets [130]. The number of these impulsive *avalanches* for a given size often forms a power law $D(s) \sim s^{-\tau}$ over many decades of sizes (Fig. 12.3(b)). In the last few decades, it has been recognized that many of these systems can also be studied as critical points—continuous transitions between qualitatively different states. We can understand most of the properties of large avalanches in these systems using the same tools developed for studying equilibrium phase transitions.

The renormalization-group and scaling methods we use to study these critical points are deep and powerful. Much of the history and practice in the field revolves around complex schemes to implement these methods for various specific systems. In this chapter, we will focus on the key ideas most useful in exploring experimental systems and new theoretical models, and will not cover the methods for calculating critical exponents.

In Section 12.1 we will examine the striking phenomenon of *universality*: two systems, microscopically completely different, can exhibit pre-

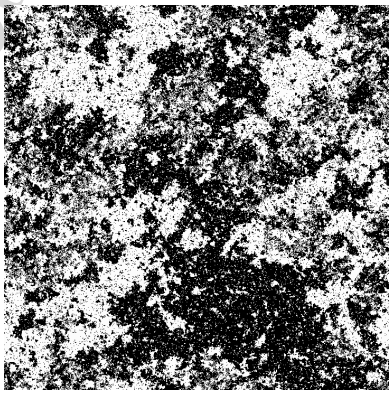


Fig. 12.1 The Ising model at T_c, the critical temperature separating the magnetized phase $T < T_c$ from the zero-magnetization phase $T > T_c$. The white and black regions represent positive and negative magnetizations $s = \pm 1$. Unlike the abrupt transitions studied in Chapter 11, here the magnetization goes to zero continuously as $T \to T_c$ from below.

[1] The term *fractal* was coined to describe sets which have characteristic dimensions that are not integers; it roughly corresponds to non-integer Hausdorff dimensions in mathematics. The term has entered the popular culture, and is associated with strange, rugged sets like those depicted in the figures here.

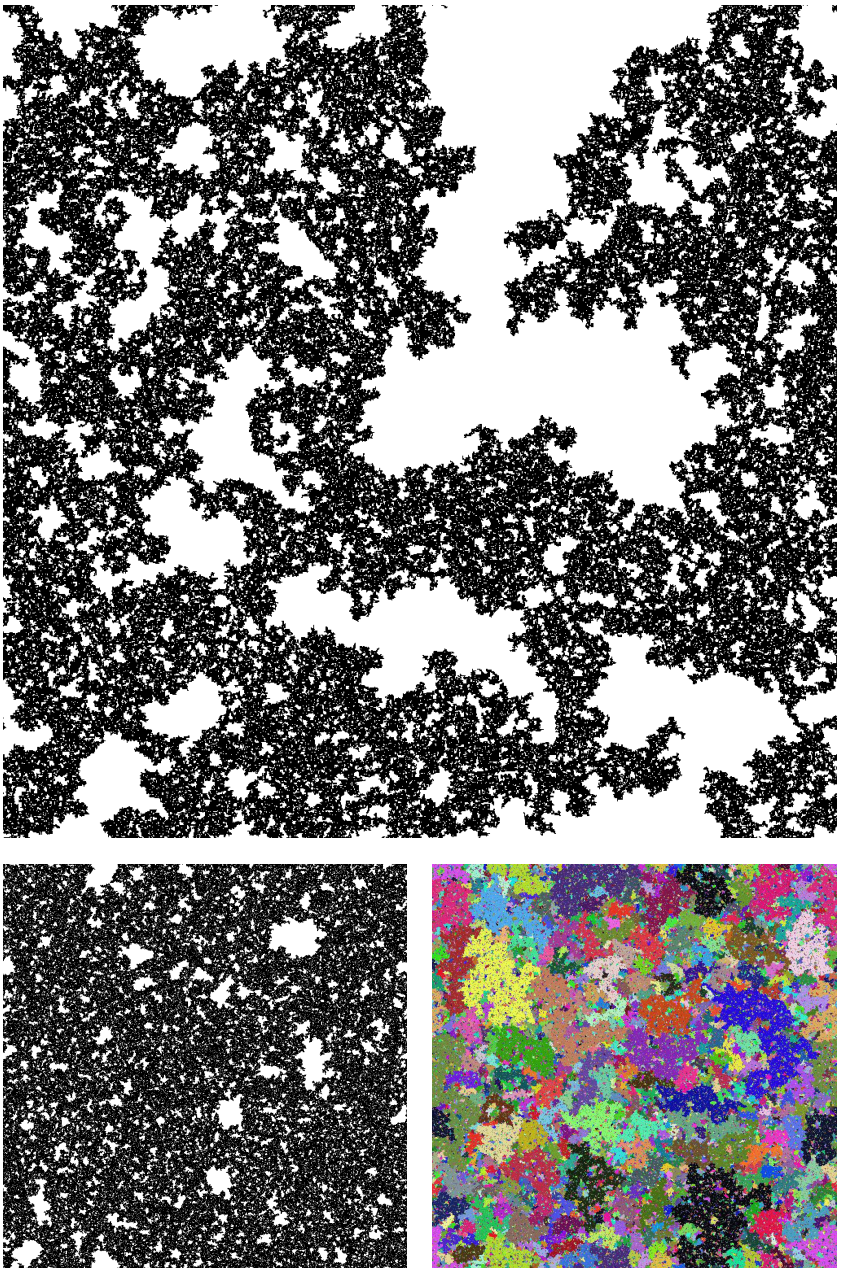


Fig. 12.2 Percolation transition. A percolation model on the computer, where bonds between grid points are removed rather than circular holes. Let the probability of removing a bond be $1-p$; then for p near one (no holes) the conductivity is large, but decreases as p decreases. After enough holes are punched (at $p_c = 1/2$ for this model), the biggest cluster just barely hangs together, with holes on all length scales. At larger probabilities of retaining bonds $p = 0.51$, the largest cluster is intact with only small holes (bottom left); at smaller $p = 0.49$ the sheet falls into small fragments (bottom right; shadings denote clusters). Percolation has a phase transition at p_c, separating a connected phase from a fragmented phase (Exercises 2.20 and 12.25).

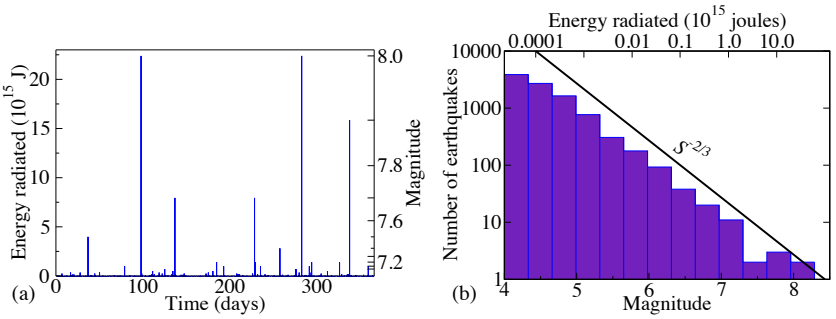


Fig. 12.3 Earthquake sizes. (a) Earthquake energy release in 1995 versus time. This time series, when sped up, sounds like crackling noise [72]. (b) Histogram of the number of earthquakes in 1995 as a function of their size S. Notice the logarithmic scales; the smallest earthquakes shown are a million times smaller and a thousand times more probable than the largest earthquakes. The fact that this distribution is well described by a power law is the Gutenberg–Richter law $\sim S^{-2/3}$.

cisely the same critical behavior near their phase transitions. We will provide a theoretical rationale for universality in terms of a *renormalization-group* flow in a space of all possible systems.

In Section 12.2 we will explore the characteristic *self-similar* structures found at continuous transitions. Self-similarity is the explanation for the fractal-like structures seen at critical points: a system at its critical point looks the same when rescaled in length (and time). We will show that *power laws* and *scaling functions* are simply explained from the assumption of self-similarity.

Finally, in Section 12.3 we will give an overview of the wide variety of types of systems that are being understood using renormalization-group and scaling methods.

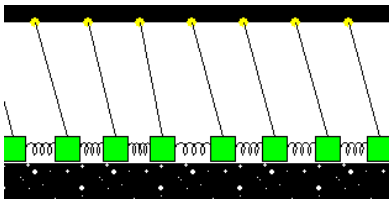


Fig. 12.4 The Burridge–Knopoff model of earthquakes, with the earthquake fault modeled by blocks pulled from above and sliding with friction on a surface below. It was later realized by Carlson and Langer [26] that this model evolves into a state with a large range of earthquake sizes even for regular arrays of identical blocks.

12.1 Universality

Quantitative theories of physics are possible because macroscale phenomena are often independent of microscopic details. We saw in Chapter 2 that the diffusion equation was largely independent of the underlying random collision processes. Fluid mechanics relies upon the emergence of simple laws—the Navier-Stokes equations—from complex underlying microscopic interactions; if the macroscopic fluid motions depended in great detail on the shapes and interactions of the constituent molecules, we could not write simple continuum laws. Ordinary quantum mechanics relies on the fact that the behavior of electrons, nuclei, and photons are largely independent of the details of how the nucleus is assembled—non-relativistic quantum mechanics is an effective theory which emerges out of more complicated unified theories at low energies. High-energy particle theorists developed the original notions of *renormalization* in order to understand how these effective theories emerge in relativistic quantum systems. Lattice quantum chromodynamics (simulating the strong interaction which assembles the nucleus) is useful only because a lattice simulation which breaks translational, rotational, and Lorentz symmetries can lead on long length scales to a behavior that nonetheless exhibits these symmetries. In each of these fields of physics,

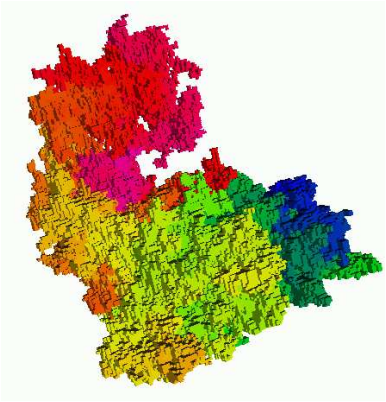


Fig. 12.5 A medium-sized avalanche (flipping 282 785 domains) in a model of avalanches and hysteresis in magnets [130] (see Exercises 8.17, 12.26 and Fig. 12.11). The shading depicts the time evolution: the avalanche started in the dark region in the back, and the last spins to flip are in the upper, front region. The sharp changes in shading are real, and represent sub-avalanches separated by times where the avalanche almost stops (see Fig. 8.19).

many different microscopic models lead to the same low-energy, long-wavelength theory.

The behavior near continuous transitions is unusually independent of the microscopic details of the system—so much so that we give a new name to it, *universality*. Figure 12.6(a) shows that the liquid and gas densities $\rho_\ell(T)$ and $\rho_g(T)$ for a variety of atoms and small molecules appear quite similar when rescaled to the same critical density and temperature. This similarity is partly for mundane reasons: the interactions between the molecules is roughly the same in the different systems up to overall scales of energy and distance. Hence argon and carbon monoxide satisfy

$$\rho^{\mathrm{CO}}(T) = A\rho^{\mathrm{Ar}}(BT) \tag{12.1}$$

for some overall changes of scale A, B. However, Fig. 12.6(b) shows a completely different physical system—interacting electronic spins in manganese fluoride, going through a ferromagnetic transition. The magnetic and liquid–gas theory curves through the data are the same if we allow ourselves to not only rescale T and the order parameter (ρ and M, respectively), but also allow ourselves to use a more general coordinate change

$$\rho^{\mathrm{Ar}}(T) = A(M(BT), T) \tag{12.2}$$

which untilts the axis.[2] Nature does not anticipate our choice of ρ and T for variables. At the liquid–gas critical point the natural measure of density is temperature dependent, and $A(M,T)$ is the coordinate change to the natural coordinates. Apart from this choice of variables, this magnet and these liquid–gas transitions all behave the same at their critical points.

This would perhaps not be a surprise if these two phase diagrams had parabolic tops; the local maximum of an analytic curve generically[3] looks parabolic. But the jumps in magnetization and density near T_c both vary as $(T_c - T)^\beta$ with the same exponent $\beta \approx 0.325$, distinctly different from the square-root singularity $\beta = 1/2$ of a generic analytic function.

Also, there are many other properties (susceptibility, specific heat, correlation lengths) which have power-law singularities at the critical point, and all of the exponents of these power laws for the liquid–gas systems agree with the corresponding exponents for the magnets. This is universality. When two different systems have the same singular properties at their critical points, we say they are in the same *universality class*. Importantly, the theoretical Ising model (despite its drastic simplification of the interactions and morphology) is also in the same universality class as these experimental uniaxial ferromagnets and liquid–gas systems—allowing theoretical physics to be directly predictive in real experiments.

To get a more clear feeling about how universality arises, consider site and bond percolation in Fig. 12.7. Here we see two microscopically different systems (left) from which basically the same behavior emerges (right) on long length scales. Just as the systems approach the threshold

[2] Here $B = T_c^M/T_c^{\ell g}$ is as usual the rescaling of temperature and $A(M,T) = a_1 M + a_2 + a_3 T = (\rho_c\rho_0/M_0)M + \rho_c(1+s) - (\rho_c s/T_c^{\ell g})T$ is a simple shear coordinate transformation from $(\rho, T^{\ell g})$ to (M, T^M). As it happens, there is another correction proportional to $(T_c - T)^{1-\alpha}$, where $\alpha \sim 0.1$ is the specific heat exponent. It can also be seen as a kind of tilt, from a pressure-dependent effective Ising-model coupling strength. It is small for the simple molecules in Fig. 12.6(a), but significant for liquid metals [49]. Both the tilt and this $1-\alpha$ correction are *subdominant*, meaning that they vanish faster as we approach T_c than the order parameter $(T_c - T)^\beta$.

[3] The term *generic* is a mathematical term which roughly translates as 'except for accidents of zero probability', like finding a function with zero second derivative at the maximum.

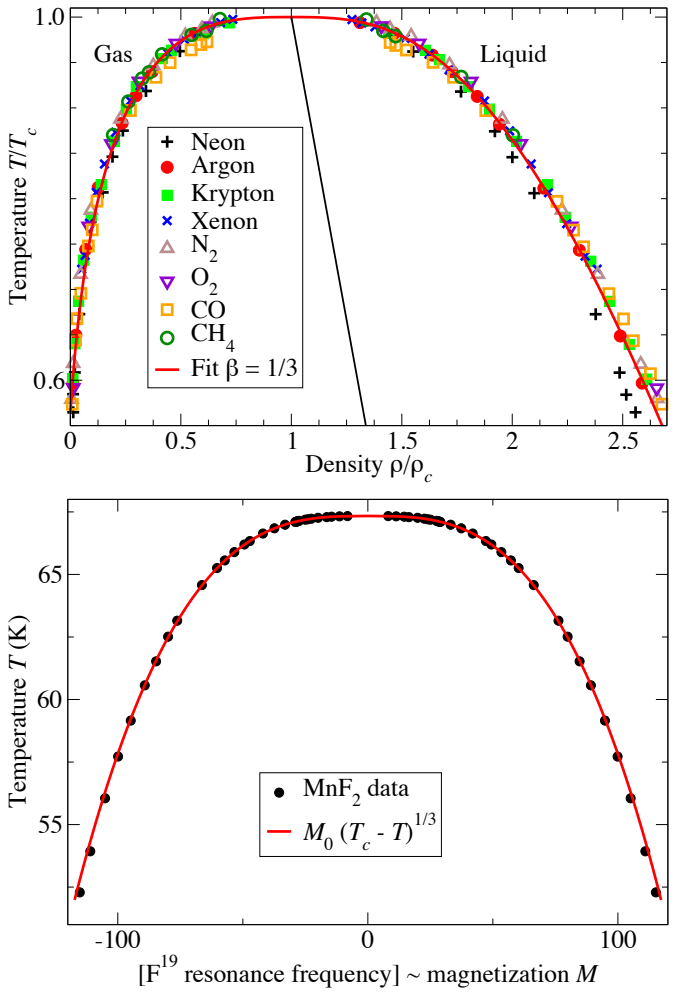


Fig. 12.6 Universality. (a) Universality at the liquid–gas critical point. The liquid–gas coexistence lines ($\rho(T)/\rho_c$ versus T/T_c) for a variety of atoms and small molecules, near their critical points (T_c, ρ_c) [56]. The curve is a fit to the argon data, $\rho/\rho_c = 1 + s(1 - T/T_c) \pm \rho_0(1 - T/T_c)^\beta$ with $s = 0.75$, $\rho_0 = 1.75$, and $\beta = 1/3$ [56]. (b) Universality: ferromagnetic–paramagnetic critical point. Magnetization versus temperature for a uniaxial antiferromagnet MnF_2 [58]. We have shown both branches $\pm M(T)$ and swapped the axes so as to make the analogy with the liquid–gas critical point (above) apparent. Notice that both the magnet and the liquid–gas critical point have order parameters that vary as $(1 - T/T_c)^\beta$ with $\beta \approx 1/3$. (The liquid–gas coexistence curves are tilted; the two theory curves would align if we defined an effective magnetization for the liquid–gas critical point $\rho_{\text{eff}} = \rho - 0.75\rho_c(1 - T/T_c)$ (thin midline, above). This is not an accident; both are in the same universality class, along with the three-dimensional Ising model, with the current estimate for $\beta = 0.325 \pm 0.005$ [152, chapter 28].

of falling apart, they become similar to one another! In particular, all signs of the original lattice structure and microscopic rules have disappeared.[4]

Thus we observe in these cases that different microscopic systems look the same near critical points, if we ignore the microscopic details and confine our attention to long length scales. To study this systematically, we need a method to take a kind of continuum limit, but in systems which remain inhomogeneous and fluctuating even on the largest scales. This systematic method is called the *renormalization group*.[5]

The renormalization group starts with a remarkable abstraction: it

[4]Notice in particular the *emergent symmetries* in the problem. The large percolation clusters at p_c are statistically both translation invariant and rotation invariant, independent of the grids that underly them. In addition, we will see that there is an emergent *scale invariance*—a kind of symmetry connecting different length scales (as we also saw for random walks, Fig. 2.2).

[5]The word renormalization grew out of quantum electrodynamics, where the effective charge on the electron changes size (norm) as a function of length scale. The word group is usually thought to refer to the family of coarse-graining operations that underly the method (with the group product being repeated coarse-graining). However, there is no inverse operation to coarse-graining, so the renormalization group does not satisfy the definition of a mathematical group.

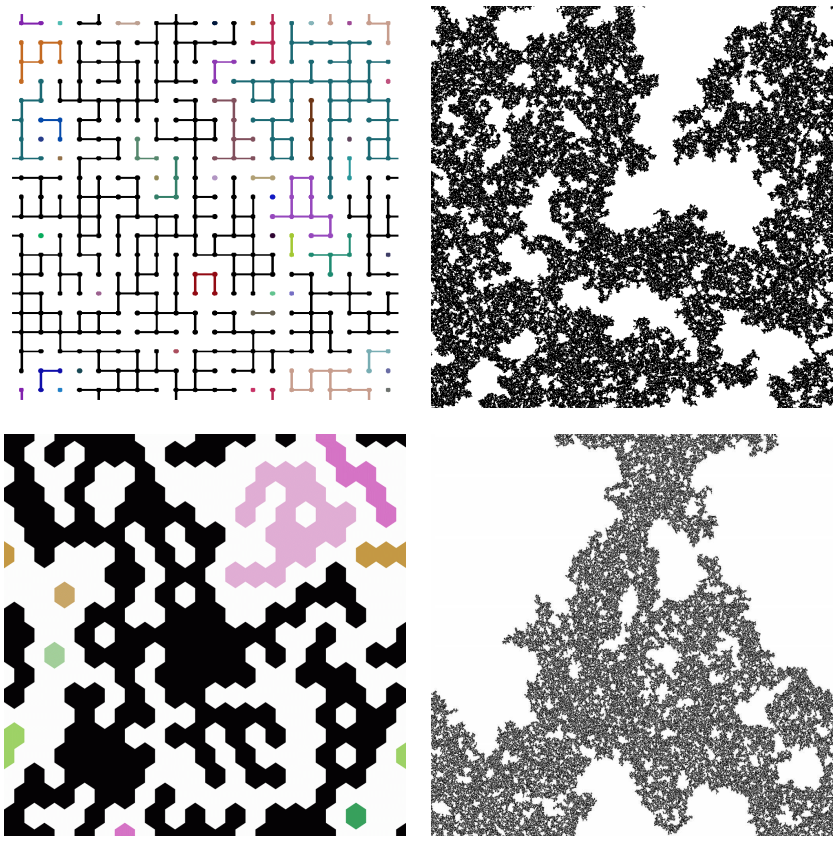


Fig. 12.7 Universality in percolation. Universality suggests that the entire morphology of the percolation cluster at p_c should be independent of microscopic details. On the top, we have bond percolation, where the bonds connecting nodes on a square lattice are occupied at random with probability p; the top right shows the infinite cluster on a 1024×1024 lattice at $p_c = 0.5$. On the bottom, we have site percolation on a triangular lattice, where it is the hexagonal sites that are occupied with probability $p = p_c = 0.5$. Even though the microscopic lattices and occupation rules are completely different, the resulting clusters look statistically identical. (One should note that the site percolation cluster is slightly less dark. Universality holds up to overall scale changes, here up to a change in the density.)

works in an enormous 'system space'. Different points in system space represent different materials under different experimental conditions, and different physical models of these materials with different interactions and evolution rules. So, for example, in Fig. 12.8 we can consider the space of all possible models for hysteresis and avalanches in three-dimensional systems. There is a different dimension in this system space for each possible parameter in a theoretical model (disorder, coupling, next-neighbor coupling, dipole fields, ...) and also for each parameter in an experiment (chemical composition, temperature, annealing time, ...). A given experiment or theoretical model will traverse a line in system space as a parameter is varied; the line at the top of the figure might represent an avalanche model (Exercise 8.17) as the strength of the disorder R is varied.

The renormalization group studies the way in which system space maps into itself under *coarse-graining*. The coarse-graining operation shrinks the system and removes microscopic degrees of freedom. Ignoring the microscopic degrees of freedom yields a new physical system with identical long-wavelength physics, but with different (renormalized) values of the parameters. As an example, Fig. 12.9 shows a

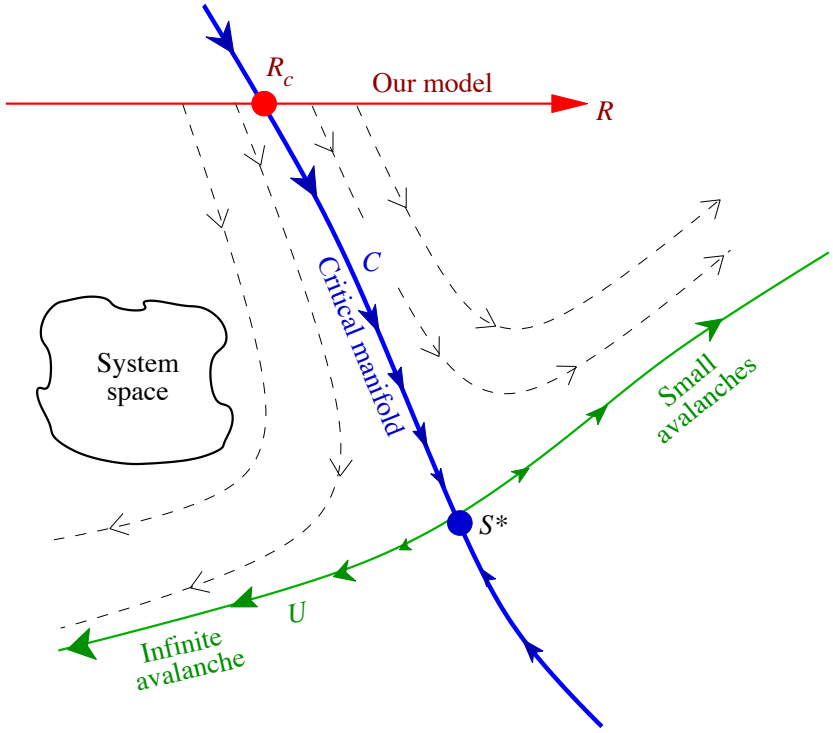


Fig. 12.8 The renormalization group defines a mapping from the space of physical systems into itself using a coarse-graining procedure. Consider the system space of all possible models of avalanches in hysteresis [130]. Each model can be coarse-grained into a new model, removing some fraction of the microscopic degrees of freedom and introducing new rules so that the remaining domains still flip at the same external fields. A fixed-point S^* under this coarse-graining mapping will be self-similar (Fig. 12.11) because it maps into itself under a change in length scale. Points like R_c that flow into S^* will also show the same self-similar behavior (except on short length scales that are coarse-grained away during the flow to S^*). Models at R_c and S^* share the same *universality class*. Systems near to their critical point coarse-grain away from S^* along the unstable curve U; hence they share universal properties too (Fig. 12.13).

real-space renormalization-group 'majority rule' coarse-graining procedure applied to the Ising model.[6] Several detailed mathematical techniques have been developed to implement this coarse-graining operation: not only real-space renormalization groups, but momentum-space ϵ-expansions, Monte Carlo renormalization groups, etc. These implementations are both approximate and technically challenging; we will not pursue them in this chapter (but see Exercises 12.22 and 12.24).

Under coarse-graining, we often find a fixed-point S^* for this mapping in system space. All the systems that flow into this fixed point under coarse-graining will share the same long-wavelength properties, and will hence be in the same universality class.

Figure 12.8 depicts the flows in system space. It is a two-dimensional picture of an infinite-dimensional space. You can think of it as a planar cross-section in system space, which we have chosen to include the line for our model and the fixed-point S^*; in this interpretation the arrows and flows denote projections, since the real flows will point somewhat out of the plane. Alternatively, you can think of it as the curved surface swept out by our model in system space as it coarse-grains, in which case you should ignore the parts of the figure below the curve U.[7]

Figure 12.8 shows the case of a fixed-point S^* that has one unstable direction, leading outward along U. Points deviating from S^* in that direction will not flow to it under coarse-graining, but rather will flow

[6] We will not discuss the methods used to generate effective interactions between the coarse-grained spins.

[7] The unstable manifold of the fixed-point.

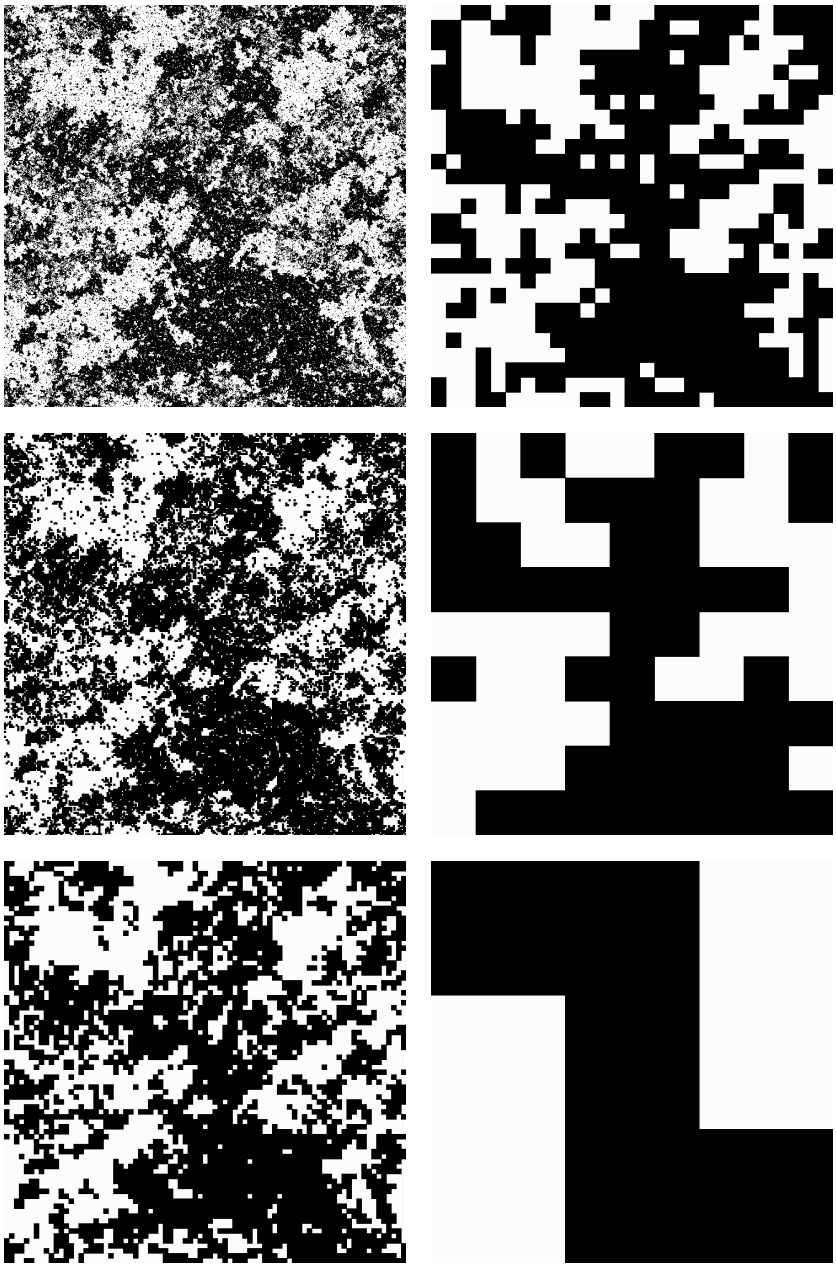


Fig. 12.9 Ising model at T_c: coarse-graining. Coarse-graining of a snapshot of the two-dimensional Ising model at its critical point. Each coarse-graining operation changes the length scale by a factor $B = 3$. Each coarse-grained spin points in the direction given by the majority of the nine fine-grained spins it replaces. This type of coarse-graining is the basic operation of the real-space renormalization group.

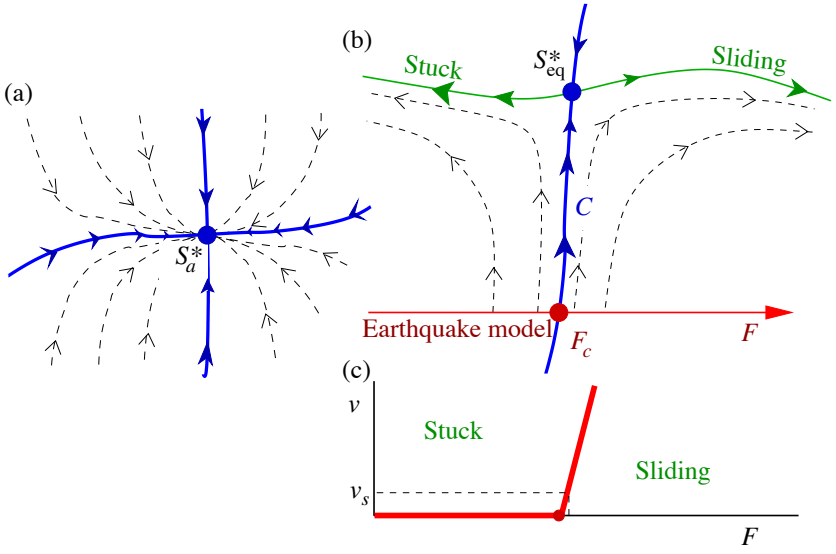


Fig. 12.10 Generic and self-organized criticality. (a) Often there will be fixed-points that attract in all directions. These fixed-points describe phases rather than phase transitions. Most phases are rather simple, with fluctuations that die away on long length scales. When fluctuations remain important, they will exhibit self-similarity and power laws called *generic scale invariance*. (b) The critical manifold C in this earthquake model separates a phase of stuck faults from a phase of sliding faults, with the transition due to the external stress F across the fault. Only along C does one find self-similar behavior and a broad spectrum of earthquakes. (c) The velocity of the fault will vary as a power law $v \sim (F - F_c)^\beta$ near the critical force F_c. The motion of the continental plates, however, drives the fault at a constant, very slow velocity v_s, automatically setting F to F_c and yielding earthquakes of all sizes; the model exhibits *self-organized criticality*.

away from it. Fixed-points with unstable directions correspond to continuous transitions between qualitatively different states. In the case of hysteresis and avalanches, there is a phase consisting of models where all the avalanches remain small, and another phase consisting of models where one large avalanche sweeps through the system, flipping most of the domains. The surface C which flows into S^* represents systems at their critical points; hence our model exhibits avalanches of all scales at R_c where it crosses C.[8]

Cases like the liquid–gas transition with two tuning parameters (T_c, P_c) determining the critical point will have fixed points with two unstable directions in system space. What happens when we have no unstable directions? The fixed-point S_a^* in Fig. 12.10 represents an entire region of system space that shares long-wavelength properties; it represents a *phase* of the system. Usually phases do not show fluctuations on all scales. Fluctuations arise near transitions because the system does not know which of the available neighboring phases to prefer. However, there are cases where the fluctuations persist even inside phases, leading to *generic scale invariance*. A good example is the case of the random walk[9] where a broad range of microscopic rules lead to the same long-wavelength random walks, and fluctuations remain important on all scales without tuning any parameters.

Sometimes the external conditions acting on a system naturally drive it to stay near or at a critical point, allowing one to spontaneously observe fluctuations on all scales. A good example is provided by certain models of earthquake fault dynamics. Fig. 12.10(b) shows the

[8] Because S^* has only one unstable direction, C has one less dimension than system space (mathematically we say C has *co-dimension* one) and hence can divide system space into two phases. Here C is the *stable manifold* for S^*.

[9] See Section 2.1 and Exercises 12.23 and 12.24.

renormalization-group flows for these earthquake models. The horizontal axis represents the external stress on the earthquake fault. For small external stresses, the faults remain stuck, and there are no earthquakes. For strong external stresses, the faults slide with an average velocity v, with some irregularities but no large events. The earthquake fixed-point $S^*_{\rm eq}$ describes the transition between the stuck and sliding phases, and shows earthquakes of all scales. The Earth, however, does not apply a constant stress to the fault; rather, continental drift applies a constant, extremely small velocity v_s (of the order of centimeters per year). Fig. 12.10(c) shows the velocity versus external force for this transition, and illustrates how forcing at a small external velocity naturally sets the earthquake model at its critical point—allowing spontaneous generation of critical fluctuations, called *self-organized criticality*.

12.2 Scale invariance

The other striking feature of continuous phase transitions is the common occurrence of self-similarity, or scale invariance. We can see this vividly in the snapshots of the critical point in the Ising model (Fig. 12.1), percolation (Fig. 12.2), and the avalanche in the hysteresis model (Fig. 12.5). Each shows roughness, irregularities, and holes on all scales at the critical point. This roughness and fractal-looking structure stems at root from a hidden symmetry in the problem: these systems are (statistically) invariant under a change in *length scale*.

Consider Figs 2.2 and 12.11, depicting the self-similarity in a random walk and a cross-section of the avalanches in the hysteresis model. In each set, the upper-left figure shows a large system, and each succeeding picture zooms in by another factor of two. In the hysteresis model, all the figures show a large avalanche spanning the system (black), with a variety of smaller avalanches of various sizes, each with the same kind of irregular boundary (Fig. 12.5). If you blur your eyes a bit, the figures should look roughly alike. This rescaling and eye-blurring process is the renormalization-group coarse-graining transformation. Figure 12.9 shows one tangible rule sometimes used to implement this coarse-graining operation, applied repeatedly to a snapshot of the Ising model at T_c. Again, the correlations and fluctuations look the same after coarse-graining; the Ising model at T_c is statistically self-similar.

This scale invariance can be thought of as an emergent symmetry under changes of length scale. In a system invariant under translations, the expectation of any function of two positions x_1, x_2 can be written in terms of the separation between the two points $\langle g(x_1, x_2)\rangle = \mathcal{G}(x_2 - x_1)$. In just the same way, scale invariance will allow us to write functions of N variables in terms of *scaling functions* of $N-1$ variables—except that these scaling functions are typically multiplied by power laws in one of the variables.

Let us begin with the case of functions of one variable. Consider the

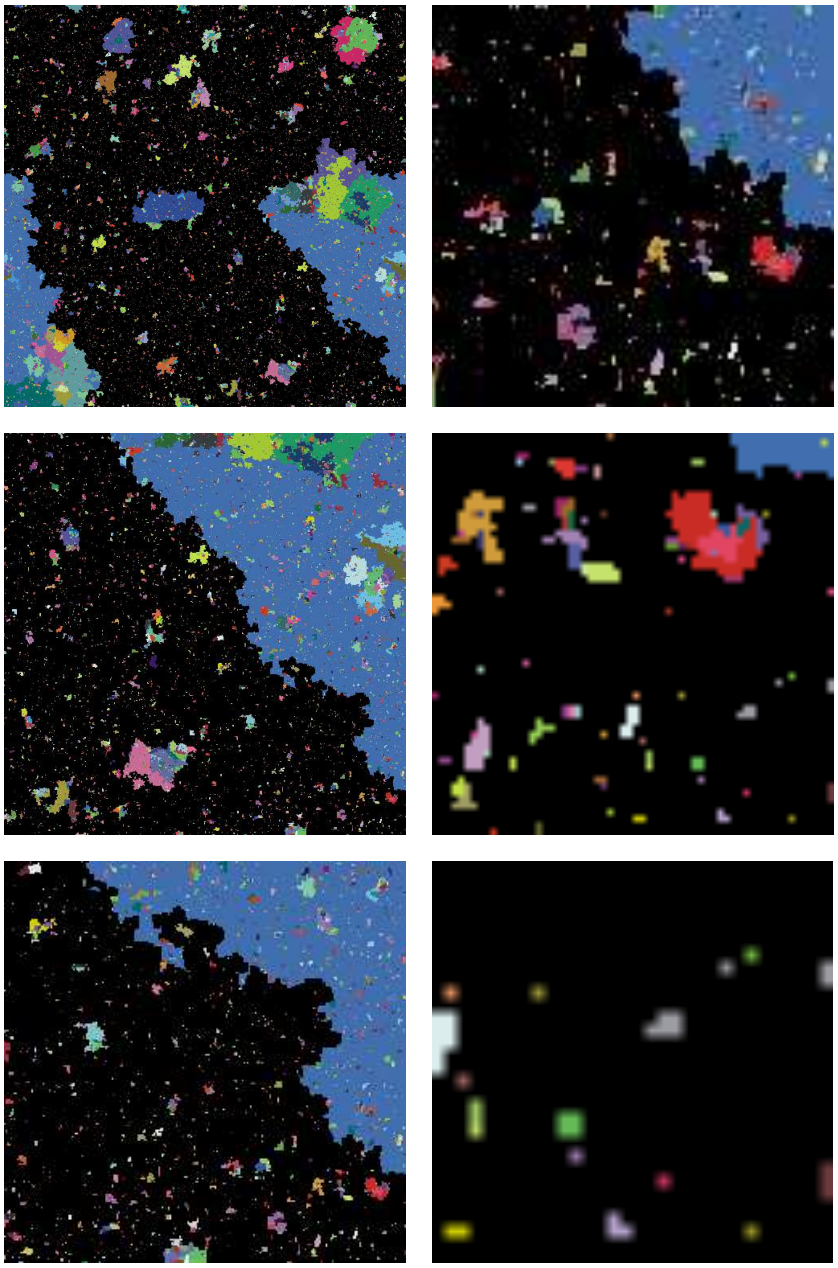


Fig. 12.11 Avalanches: scale invariance. Magnifications of a cross-section of all the avalanches in a run of our hysteresis model (Exercises 8.17 and 12.26) each one the lower right-hand quarter of the previous. The system started with a billion domains (1000^3). Each avalanche is shown in a different shade. Again, the larger scales look statistically the same.

avalanche size distribution $D(S)$ for a model, say the real earthquakes in Fig. 12.3(a), or our model for hysteresis, at the critical point. Imagine taking the same system, but increasing the units of length with which we measure the system—stepping back, blurring our eyes, and looking at the system on a coarse-grained level. Imagine that we multiply the spacing between markings on our rulers by a small amount $B = 1 + \mathrm{d}\ell$. After coarsening, any length scales in the problem (like the correlation length ξ) will be divided by B.

$$\begin{aligned} \xi' = \xi/B = \xi/(1+\mathrm{d}\ell) = \xi - \xi\mathrm{d}\ell + O(\mathrm{d}\ell^2), \\ \mathrm{d}\xi/\mathrm{d}\ell = -\xi, \\ \xi[\ell] = \xi_\ell = \xi_0 \exp(-\ell). \end{aligned} \tag{12.3}$$

(We shall denote X_ℓ the value of a quantity X after coarse-graining by $\exp(\ell)$. Thus X_0 is the initial condition to our renormalization-group flow equations, and hence is the 'true' value in the simulation or experiment.)

Thus $\exp(\ell)$ is the net coarse-graining length as a function of the parameter ℓ. We assume that the system is flowing to a fixed point under our renormalization group, so for large avalanches we expect

$$D'(S) = D(S); \qquad \partial D/\partial \ell = 0. \tag{12.4}$$

How do D and S change under the renormalization group? The avalanche sizes S after coarse-graining will be smaller by some factor[10] $C = 1 + c\mathrm{d}\ell$. The overall scale of $D(S)$ will change by some factor $A = 1 + a\mathrm{d}\ell$. This factor A is partly due to coarse graining; the same avalanches occur independent of the units of length with which we measure, but the probability density $D(S)$ per unit size per unit volume will change. It is partly due to the rescaling factor allowed by the renormalization group – here D' must omit the smallest avalanches $S < S_{\mathrm{smallest}}$ (now invisible after blurring our eyes), so the overall normalization factor dividing D changes as well under rescaling. Thus our renormalization-group equations are

$$\begin{aligned} S' = S/C = S/(1+c\mathrm{d}\ell), \qquad & \mathrm{d}S/\mathrm{d}\ell = -cS, \\ D'(S') = AD(S) = D(S)(1+a\mathrm{d}\ell), \quad & \mathrm{d}D/\mathrm{d}\ell = aD. \end{aligned} \tag{12.5}$$

Here $D'(S')$ is the distribution measured with the new ruler: a smaller avalanche with a larger probability density. Note that the flow equation 12.5 for D specifies the total derivative $\mathrm{d}D/\mathrm{d}\ell = -aD$ (see Section 4.1), while being at a fixed point (eqn 12.4). specifies the partial derivative $\partial D/\partial \ell = 0$. The total derivative is $\mathrm{d}D_\ell(S_\ell)/\mathrm{d}\ell$; it gives the change of $D'(S') - D(S)$, where the partial derivative $\partial D/\partial \ell$ gives the change $D'(S) - D(S)$.

Solving eqn 12.5 for the flow of S, we find $S_\ell = S_0 \exp(-c\ell)$. Here S_0 is both the initial condition for our differential equation at $\ell = 0$, and is the physical avalanche size for which we want to know the probability $D_0(S_0)$. Similarly, $D_\ell(S_\ell) = \exp(a\ell)D_0(S_0)$. We want to solve

[10] If the size of the avalanche were the cube of its length, then c would equal 3 since $(1 + \mathrm{d}\ell)^3 = 1 + 3\mathrm{d}\ell + O(\mathrm{d}\ell^2)$. Here c is the fractal dimension of the avalanche.

for the probability $D_0(S_0)$ in terms of S_0. We can get rid of the exponential factor by noticing $\exp(a\ell) = \exp(-c\ell)^{-a/c} = (S_\ell/S_0)^{-a/c}$, so $D_\ell(S_\ell) = (S_\ell/S_0)^{-a/c} D_0(S_0)$. We can then choose to flow until ℓ^* such that $S_{\ell^*} = 1$ (along the $r = 0$ axis in Fig. 12.12). Thus we find $D_0(S_0) = D_{\ell^*}(1) S_0^{-a/c}$, or

$$D(S) = N S^{-a/c}. \tag{12.6}$$

where N is the constant $D_{\ell^*}(1)$. This argument is quite general. Not only do we expect that avalanches (and earthquakes, Fig. 12.3) will show power laws at critical points, but almost anything which rescales by a constant factor under the renormalization group should exhibit a power law.

Because the properties shared in a universality class only hold up to overall scales, the constant N is system dependent. (In this case, $\int_{S_{\text{smallest}}}^{\infty} D(S) \mathrm{d}S = 1$ because D is a probability distribution, so the normalization factor $N = (\tau - 1) S_{\text{smallest}}^{(1-\tau)}$.) However, the exponents a, c, and a/c are *universal*—independent of experiment (with the universality class). Some of these exponents have standard names: the exponent c giving the fractal dimension of the avalanche is usually called d_f or $1/\sigma\nu$. The exponent a/c giving the size distribution law is called τ in percolation and in most models of avalanches in magnets[11] and is related to the Gutenberg–Richter exponent for earthquakes[12] (Fig. 12.3(b)).

Most measured quantities depending on one variable will have similar power-law singularities at the critical point. Thus the correlation function of the Ising model at T_c (Fig. 10.4) decays with distance x in dimension d as $C(x) \propto x^{-(d-2+\eta)}$ and the distance versus time for random walks (Section 2.1) grows as $t^{1/2}$, both because these systems are self-similar.[13]

Self-similarity is also expected near to the critical point. Here as one coarsens the length scale a system will be statistically similar to itself *at a different set of parameters*. Thus a system undergoing phase separation (Section 11.4.1, Exercise 12.17), when coarsened, is similar to itself at an earlier time (when the domains were smaller), and a percolation cluster just above p_c (Fig. 12.2 (bottom left)) when coarsened is similar to one generated further from p_c (hence with smaller holes).

For a magnet slightly below[14] T_c, a system coarsened by a factor $B = 1 + \mathrm{d}\ell$ will be similar to one farther from T_c by a factor $E = 1 + e\mathrm{d}\ell$. Here the standard Greek letter for the length rescaling exponent is $\nu = 1/e$. Similar to the case of the avalanche size distribution, the coarsened system must have its magnetization rescaled upward by $F = (1 + f\mathrm{d}\ell)$, with $f = \beta/\nu$) to match that of the lower-temperature original magnet (Fig. 12.13):

$$\begin{aligned} M'(T_c - t) &= F M(T_c - t) = M(T_c - Et), \\ (1 + f\mathrm{d}\ell)\, M(T_c - t) &= M\Big(T_c - t(1 + e\mathrm{d}\ell)\Big), \end{aligned} \tag{12.7}$$

so

$$\begin{aligned} \mathrm{d}t/\mathrm{d}\ell &= et, \\ \mathrm{d}M/\mathrm{d}\ell &= fM. \end{aligned} \tag{12.8}$$

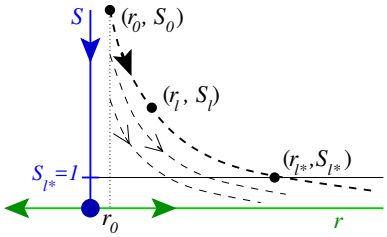


Fig. 12.12 Disorder and avalanche size: renormalization-group flows. Here S is the avalanche size and $r = R - R_c$ is the change in the disorder from the critical point. The disorder grows under coarse-graining, and the avalanche sizes shrink. The renormalization group allows us to describe properties of avalanches at any large size S and small r (near the critical disorder) as rescaled versions of the avalanches along a single line (for convenience, along $S = 1$). First, the size distribution $D(S, R_c)$ at criticality can be written in terms of the constant $N = D(1, R_c)$ rescaled by a power law (eqn 12.6) (since all points on the blue axis flow there). Later, we show that $D(S, R_c + r_0)$ can be written as the same power law times $D(1, R_c + r_{\ell^*})$, where $r_{\ell^*}(S_0)$ is the intersection of the flow trajectory starting at (r_0, S_0).

[11] Except ours, where we used τ to denote the avalanche size law at the critical field and disorder; integrated over the hysteresis loop $D_{\text{int}} \propto S^{-\bar{\tau}}$ with $\bar{\tau} = \tau + \sigma\beta\delta$.

[12] We must not pretend that we have found the final explanation for the Gutenberg–Richter law. There are many different models that give exponents $\approx 2/3$, but it remains controversial which of these, if any, are correct for real-world earthquakes.

[13] This is because power laws are the only self-similar function. If $f(x) = x^{-\alpha}$, then on a new scale multiplying x by B, $f(Bx) = B^{-\alpha} x^{-\alpha} \propto f(x)$. (See [102] for more on power laws.)

[14] Thus increasing the distance t to T_c decreases the temperature T.

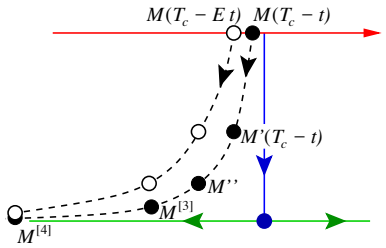


Fig. 12.13 Scaling near criticality. If two points in system space flow towards one another under coarse-graining, their behavior must be similar on long length scales. Here we measure the magnetization $M(T)$ for our system (top line) at two different temperatures, $T_c - t$ and $T_c - Et$. The dots represent successive coarse-grainings by a factor B; under this renormalization group $M \to M' \to M'' \to M^{[3]} \ldots$. Here $M(T_c - t)$ after four coarse-grainings maps to nearly the same system as $M(T_c - Et)$ after three coarse-grainings. We thus know, on long length scales, that $M'(T_c - t)$ must agree with $M(T_c - Et)$; the system is similar to itself *at a different set of external parameters*. All systems near to criticality first are attracted near to the fixed point, and then flow away along a common trajectory (here the horizontal axis, in Fig. 12.8 the unstable curve U). Their properties are universal because they all escape along the same path.

[15] They can be derived from the eigenvalues of the linearization of the renormalization-group flow around the fixed-point S^* in Fig. 12.8 (see Exercises 12.1 and 12.24).

Again, $M \propto t^{f/e} = t^{\beta}$, providing a rationale for the power laws we saw in magnetism and the liquid–gas transition (Fig. 12.6). Similarly, the specific heat, correlation length, correlation time, susceptibility, and surface tension of an equilibrium system will have power-law divergences $(T - T_c)^{-X}$, where by definition X is α, ν, $z\nu$, γ, and -2ν, respectively. One can also vary the field H away from the critical point and measure the resulting magnetization, which varies as $H^{1/\delta}$.

To specialists in critical phenomena, these exponents are central; whole conversations often rotate around various combinations of Greek letters. We know how to calculate critical exponents from the various analytical approaches,[15] and they are simple to measure (although hard to measure well, [88]).

Critical exponents are not everything, however. Many other scaling predictions are easily extracted from numerical simulations. Universality should extend even to those properties that we have not been able to write formulæ for. In particular, there are an abundance of functions of two and more variables that one can measure. Figure 12.14 shows the distribution of avalanche sizes $D_{\mathrm{int}}(S, R)$ in our model of hysteresis, integrated over the hysteresis loop (Fig. 8.16), at various disorders R above R_c (Exercise 8.17). Notice that only at $R_c \approx 2.16$ do we get a power-law distribution of avalanche sizes; at larger disorders there are extra small avalanches, and a strong decrease in the number of avalanches beyond a certain size $S_{\max}(R)$.

Let us derive the scaling form for $D_{\mathrm{int}}(S, R)$. By using scale invariance, we will be able to write this function of two variables as a power of one of the variables times a universal, one-variable function of a combined *invariant scaling combination*. As in Fig. 12.8 we expect that systems will flow away from criticality: a system at $R = R_c + r$ after coarse-graining will be similar to a system further from the critical disorder, say at $R' = R_c + Er = R_c + (1 + e\mathrm{d}\ell)r$. Together with our treatment of the avalanche sizes at R_c (eqns 12.5) we know that

$$\begin{aligned} \mathrm{d}r/\mathrm{d}\ell &= er, \\ \mathrm{d}S/\mathrm{d}\ell &= -cS, \\ \mathrm{d}D/\mathrm{d}\ell &= aD \end{aligned} \tag{12.9}$$

First, how does the deviation of the disorder r vary away from criticality? Solving $\mathrm{d}r/\mathrm{d}\ell = er$ we find $r_\ell = r_0 \exp(e\ell)$. Similarly $S_\ell = S_0 \exp(-c\ell)$, so we may write r_ℓ in terms of S_0 as $r_\ell = r_0 (S_\ell/S_0)^{-e/c}$. To derive the form for the avalanche size distribution at size S_0, we shall again renormalize until the value ℓ^* such that $S_{\ell^*} = 1$ (Fig. 12.12). Hence

$$r_{\ell^*} = r_0 S_0^{e/c} = r_0 S_0^{\sigma} \tag{12.10}$$

where $1/\sigma = c/e$ is the exponent governing how the cutoff in the avalanche size distribution varies with disorder (inset, Fig. 12.14).

The combination $X = rS^{\sigma}$ in eqn 12.10 is *invariant* under the renormalization group flow. That is, each dashed curve in Fig 12.12 corresponds to a different value of this invariant scaling combination. It

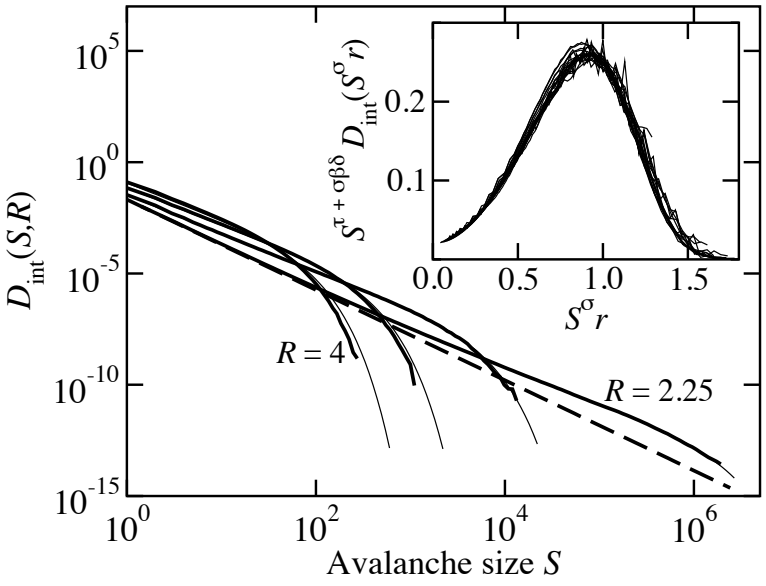


Fig. 12.14 Avalanche size distribution. The distribution of avalanche sizes in our model for hysteresis. Notice the logarithmic scales. (We can measure a $D(S)$ value of 10^{-14} by running billions of spins and binning over ranges $\Delta S \sim 10^5$.) (i) Although only at $R_c \approx 2.16$ do we get a pure power law (dashed line, $D(S) \propto S^{-\bar{\tau}}$), we have large avalanches with hundreds of spins even a factor of two away from the critical point. (ii) The curves have the wrong slope except very close to the critical point; be warned that a power law over two decades (although often publishable [88]) may not yield a reliable exponent. (iii) The scaling curves (thin lines) work well even far from R_c. Inset: We plot $D(S)/S^{-\bar{\tau}}$ versus $S^\sigma(R - R_c)/R$ to extract the universal scaling curve $\mathcal{D}(X)$ (eqn 12.11). Varying the critical exponents and R_c to get a good collapse allows us to measure the exponents far from R_c, where power-law fits are still unreliable (Exercise 12.25(g)).

makes sense that such invariants would be important. The avalanches of size S_0 in a system with disorder shifted by r_0 will have properties (duration, or their average shape, or the probability $D_0(S_0, r_0)$ we calculate here), which are similar to other points on the same curve.

Now, how does D renormalize? Solving $\mathrm{d}D/\mathrm{d}\ell = aD$, $D_\ell(S_\ell, r_\ell) = \exp(a\ell)D_0(S_0, r_0) = (S_\ell/S_0)^{-a/c}D_0(S_0, r_0)$, so $D_0 = S_0^{-a/c}D_{\ell^*}(1, r_{\ell^*})$. Using $a/c = \bar{\tau}$ and eqn 12.10, we find for the 'physical' variables $D = D_0$, $S = S_0$, and $r = r_0$ that

$$D(S, R) = S^{-\bar{\tau}}\mathcal{D}(rS^\sigma) = S^{-\bar{\tau}}\mathcal{D}((R - R_c)S^\sigma) \tag{12.11}$$

for the *scaling function* $\mathcal{D}(X) = D_{\ell^*}(1, X)$. This scaling function is another universal prediction of the theory (up to an overall choice of units for S, r and D). In Fig. 12.12, we have expressed the avalanche size distribution $D(S, R_c + r)$ for any point in the (r, S) plane in terms of the values $D(1, R_c + rS^\sigma)$. The emergent scale invariance allows us to write a function of two variables in terms of a universal function of one variable.

We can use a *scaling collapse* of the experimental or numerical data to extract this universal function, by plotting $D/S^{-\bar{\tau}}$ against $X = S^\sigma(R - R_c)$; the inset of Fig. 12.14 shows this scaling collapse.

Similar universal scaling functions appear in many contexts. Considering just the equilibrium Ising model, there are scaling functions for the magnetization $M(H, T) = (T_c - T)^\beta \mathcal{M}\left(H/(T_c - T)^{\beta\delta}\right)$, for the correlation function $C(x, t, T) = x^{-(d-2+\eta)}\mathcal{C}\left(x/|T - T_c|^{-\nu}, t/|T - T_c|^{-z\nu}\right)$, and for finite-size effects $M(T, L) = (T_c - T)^\beta \mathcal{M}\left(L/(T_c - T)^{-\nu}\right)$ in a

system confined to a box of size L^d.

12.3 Examples of critical points

Ideas from statistical mechanics have found broad applicability in sciences and intellectual endeavors far from their roots in equilibrium thermal systems. The scaling and renormalization-group methods introduced in this chapter have seen a particularly broad range of applications; we will touch upon a few in this conclusion to our text.

12.3.1 Equilibrium criticality: energy versus entropy

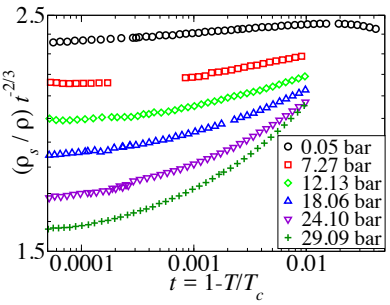


Fig. 12.15 Superfluid density in helium: scaling plot. This classic experiment [1,55] in 1980 measured the superfluid density $\rho_s(T)$ in helium to great precision. Notice the logarithmic scale on the horizontal axis; the lowest pressure data (saturated vapor pressure ≈ 0.0504 bar) spans three decades of temperature shift from T_c. This plot emphasizes the deviations from the expected power law.

[16]Potts models are Ising-like models with N states per site rather than two.

Scaling and renormalization-group methods have their roots in the study of continuous phase transitions in equilibrium systems. Ising models, Potts models,[16] Heisenberg models, phase transitions in liquid crystals, wetting transitions, equilibrium crystal shapes (Fig. 11.6), two-dimensional melting—these are the grindstones on which our renormalization-group tools were sharpened.

The transition in all of these systems represents the competition between energy and entropy, with energy favoring order at low temperatures and entropy destroying it at high temperatures. Figure 12.15 shows the results of a classic, amazing experiment—the analysis of the superfluid transition in helium (the same order parameter, and also the same universality class, as the XY model). The superfluid density is expected to have the form

$$\rho_s \propto (T_c - T)^\beta (1 + d(T_c - T)^x), \tag{12.12}$$

where x is a universal, subdominant *correction to scaling*. Since $\beta \approx \frac{2}{3}$, they plot $\rho_s/(T - T_c)^{2/3}$ so that deviations from the simple expectation are highlighted. The slope in the top, roughly straight curve reflects the difference between their measured value of $\beta = 0.6749 \pm 0.0007$ and their multiplier $\frac{2}{3}$. The other curves show the effects of the subdominant correction, whose magnitude d increases with increasing pressure. Recent experiments improving on these results were done on the space station, in order to reduce the effects of gravity.

12.3.2 Quantum criticality: zero-point fluctuations versus energy

Thermal fluctuations do not exist at zero temperature, but there are many well-studied *quantum phase transitions* which arise from the competition of potential energy and quantum fluctuations. Many of the earliest studies focused on the metal–insulator transition and the phenomenon of *localization*, where disorder can lead to insulators even when there are states at the Fermi surface. Scaling and renormalization-group methods played a central role in this early work; for example, the states

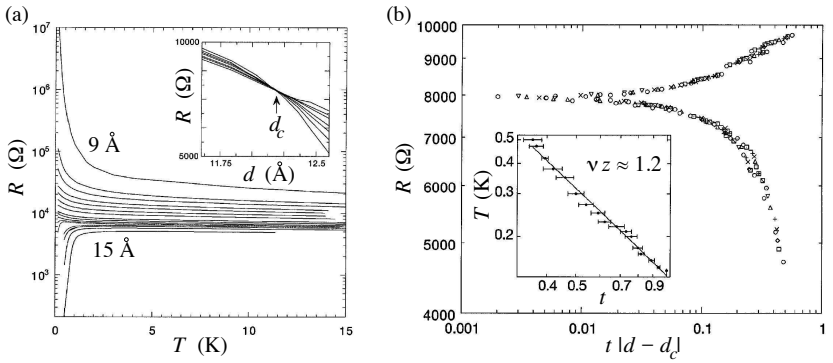


Fig. 12.16 The superconductor–insulator transition. (a) Thin films of amorphous bismuth are insulators (resistance grows to infinity at zero temperature), while films above about 12 Å are superconducting (resistance goes to zero at a temperature above zero). (b) Scaling collapse. Resistance plotted against the scaled thickness for the superconductor–insulator transition, with each thickness rescaled by an independent factor t to get a good collapse. The top scaling curve $\mathcal{F}_-$ is for the insulators $d < d_c$, and the bottom one $\mathcal{F}_+$ is for the superconductors $d > d_c$. The inset shows $t \sim T^{-1/\nu z}$, with $\nu z \sim 1.2$. (From [91].)

near the *mobility edge* (separating localized from extended states) are self-similar and fractal. Other milestones include the Kondo effect, macroscopic quantum coherence (testing the fundamentals of quantum measurement theory), transitions between quantum Hall plateaus, and superconductor–normal metal transitions. Figure 12.16 show a recent experiment studying a transition directly from a superconductor to an insulator, as the thickness of a film is varied. The resistance is expected to have the scaling form

$$R(d,T) = R_c \mathcal{F}_\pm\Big((d-d_c)T^{-1/\nu z}\Big); \tag{12.13}$$

the authors plot $R(d,T)$ versus $t(d-d_c)$, vary t until the curves collapse (main part of Fig. 12.16(b)), and read off $1/\nu z$ from the plot of t versus T (inset). While it is clear that scaling and renormalization-group ideas are applicable to this problem, we should note that as of the time this text was written, no theory yet convincingly explains these particular observations.

12.3.3 Dynamical systems and the onset of chaos

Much of statistical mechanics focuses on systems with large numbers of particles, or systems connected to a large external environment. Continuous transitions also arise in isolated or simply driven systems with only a few important degrees of freedom, where they are called *bifurcations*. A bifurcation is a qualitative change in behavior which arises when a parameter in a set of differential equations passes through a critical value. The study of these bifurcations is the theory of *normal forms* (Exercise 12.18). Bifurcation theory contains analogies to universality classes, critical exponents, and analytic corrections to scaling.

Dynamical systems, even when they contain only a few degrees of freedom, can exhibit immensely complex, chaotic behavior. The mathematical trajectories formed by chaotic systems at late times—the *attractors*—are often fractal in structure, and many concepts and methods from

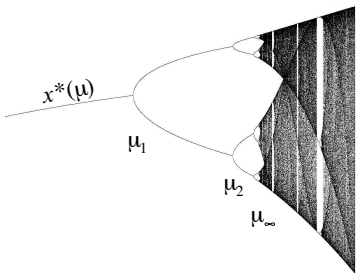


Fig. 12.17 Self-similarity at the onset of chaos. The attractor as a function of μ for the Feigenbaum logistic map $f(x) = 4\mu x(1-x)$. For small $\mu < \mu_1$, repeatedly iterating f converges to a fixed-point $x^*(\mu)$. As μ is raised past μ_1, the map converges into a two-cycle; then a four-cycle at μ_2, an eight-cycle at $\mu_3 \ldots$ These period-doubling bifurcations converge geometrically: $\mu_\infty - \mu_n \propto \delta^{-n}$ where $\delta = 4.669201609102990\ldots$ is a universal constant. At μ_∞ the system goes chaotic. (Exercise 12.22).

[17]For example, statistical mechanical ensembles become *invariant measures* (Exercise 4.10), and the attractors are characterized using concepts related to entropy (Exercise 5.23).

statistical mechanics are useful in studying these sets.[17]

It is in the study of the onset of chaos where renormalization-group methods have had a spectacular impact. Figure 12.17 shows a simple dynamical system undergoing a series of bifurcations leading to a chaotic state. Feigenbaum (Exercise 12.22) analyzed the series using a renormalization group, coarse-graining not in space but in *time*. Again, this behavior is universal—exactly the same series of bifurcations (up to smooth coordinate changes) arise in other maps and in real physical systems. Other renormalization-group calculations have been important for the study of the transition to chaos from quasiperiodic motion, and for the breakdown of the last non-chaotic region in Hamiltonian systems (see Exercise 4.7).

12.3.4 Glassy systems: random but frozen

Let us conclude with a common continuous transition for which our understanding remains incomplete: glass transitions.

Glasses are out of equilibrium; their relaxation times diverge as they are cooled, and they stop rearranging at a typical temperature known as the glass transition temperature. Many other disordered systems also appear to be glassy, in that their relaxation times get very slow as they are cooled, and they freeze into disordered configurations.[18] This freezing process is sometimes described as developing *long-range order in time*, or as a *broken ergodicity* (see Section 4.2).

[18]Glasses are different from disordered systems. The randomness in disordered systems is fixed, and occurs in both the high- and low-temperature phases; the disorder in the traditional *configurational* glasses freezes in as it cools. See also Section 5.2.2.

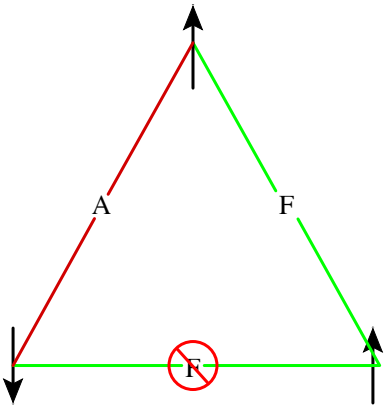


Fig. 12.18 Frustration. A spin glass has a collection of magnetic ions with interactions of random sign. Here we see a triangle of Ising ±1 spins with one antiferromagnetic bond—one of the three bonds must be unsatisfied in any spin configuration. Hence the system is said to be *frustrated*.

The basic reason that many of the glassy systems freeze into random states is *frustration*. Frustration was defined first for spin glasses, which are formed by randomly substituting magnetic atoms into a nonmagnetic host. The magnetic spins are coupled to one another at random; some pairs prefer to be parallel (ferromagnetic couplings) and some antiparallel (antiferromagnetic). Whenever strongly-interacting spins form a loop with an odd number of antiferromagnetic bonds (Fig. 12.18) they are frustrated; one of the bonds will have to be left in an unhappy state, since there must be an even number of spin inversions around the loop (Fig. 12.18). It is believed in many cases that frustration is also important for configurational glasses (Fig. 12.19).

The study of disordered magnetic systems is mathematically and computationally sophisticated. The equilibrium ground state for the three-dimensional random-field Ising model,[19] for example, has been rigorously proven to be ferromagnetic (boring); nonetheless, when cooled in zero external field we understand why it freezes into a disordered state, because the coarsening process develops diverging free energy barriers to relaxation. Methods developed to study the spin glass transition have seen important applications in neural networks (which show a forgetting transition as the memory becomes overloaded) and more recently in guiding algorithms for solving computationally hard (**NP**–complete) problems (see Exercises 1.10 and 8.19). Some basic conceptual questions, however, remain unanswered. For example, we still do not know whether spin glasses have a finite or infinite number of equilibrium states—whether,

[19]Our model for hysteresis and avalanches (Figs 8.19, 12.5, 12.11, and 12.14; Exercises 8.17, 8.18, and 12.26) is this same random-field Ising model, but in a growing external field and out of equilibrium.

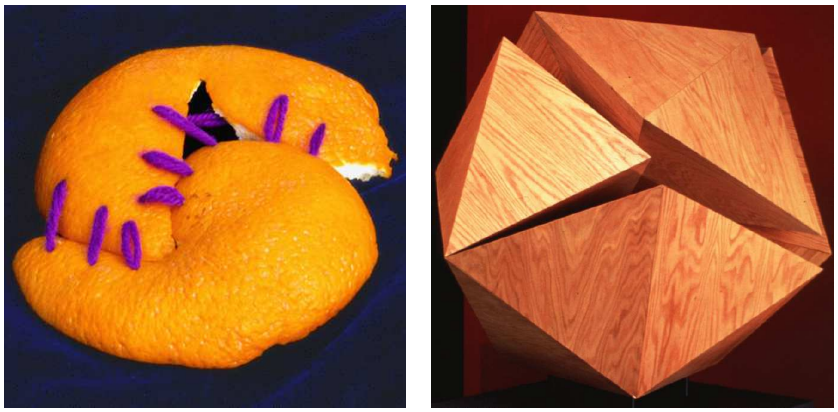


Fig. 12.19 Frustration and curvature. One kind of frustration arises when the energetically favorable local packing of atoms or molecules is incompatible with the demands of building large structures. Here we show two artistic renditions (courtesy of Pamela Davis Kivelson [70]). (a) The classic problem faced by map-makers: the peel of an orange (or the crust of the Earth) cannot be mapped smoothly onto a flat space without stretching and tearing it. (b) The analogous problem faced in many metallic glasses, whose atoms locally prefer to form nice compact tetrahedra: twenty tetrahedra cannot be glued together to form an icosahedron. Just as the orange peel segments can be nicely fit together on the sphere, the metallic glasses are unfrustrated in curved space [120].

upon infinitely slow cooling, one still has many glassy configurations.[20]

In real configurational glasses the viscosity and relaxation times grow by ten to fifteen orders of magnitude in a relatively small temperature range, until the cooling rate out-paces the equilibration. We fundamentally do not know why the viscosity diverges so rapidly in so many materials. There are at least three competing pictures for the glass transition. (1) It reflects an underlying equilibrium transition to an ideal, zero-entropy glass state, which would be formed under infinitely slow cooling. (2) It is a purely dynamical transition (where the atoms or molecules jam together). (3) It is not a transition at all, but just a cross-over where the liquid viscosity jumps rapidly (say, because of the formation of semipermanent covalent bonds).

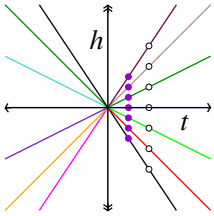
[20] There are 'cluster' theories which assume two (spin-flipped) ground states, competing with 'replica' and 'cavity' methods applied to infinite-range models which suggest many competing ground states. Some rigorous results are known.

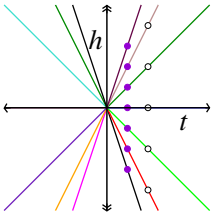
12.3.5 Perspectives

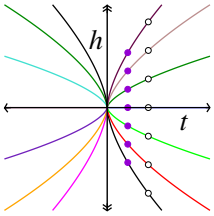
Many of the physicists who read this text will spend their careers outside of traditional physics. Physicists continue to play significant roles in the financial world (econophysics, computational finance, derivative trading), in biology (bioinformatics, models of ecology and evolution), computer science (traffic models, algorithms for solving hard problems), and to some extent in social science modeling (models of voting behavior and consensus building). The tools and methods of statistical mechanics (particularly the scaling methods used to study continuous transitions) are perhaps the most useful tools that we bring to these disparate subjects. Conversely, I hope that this text will prove useful as an introduction of these tools and methods to computer scientists, biologists, engineers, and finance professionals, as they continue to broaden and fertilize the field of statistical mechanics.

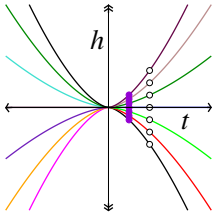
Exercises

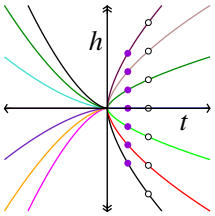
We start by re-emphasizing the phenomena and conventions used in studying continuous phase transitions in *Ising self-similarity*, *Scaling functions*, and *Scaling and coarsening*. We study solvable critical points in *Bifurcation theory*, *Mean-field theory* and *The onset of lasing*. We illustrate how the renormalization-group flows determine the critical behavior in *Renormalization-group trajectories* and *Superconductivity and the renormalization group*; the latter explains schematically the fundamental basis for Fermi liquid theory. *Period doubling* and the two versions of *The renormalization-group and the central limit theorem* provide important applications where the reader may implement the renormalization group explicitly and completely. We conclude with two numerical exercises, *Percolation and universality* and *Hysteresis and avalanches: scaling*, which mimic the entire experimental analysis from data collection to critical exponents and scaling functions.

(A) 

(B) 

(C) 

(D) 

(E) 

(12.1) **Renormalization-group trajectories.** ③
This exercise provides an early introduction to how we will derive power laws and universal scaling functions in Section 12.2 from universality and coarse-graining.
An Ising model near its critical temperature T_c is described by two variables: the distance to the critical temperature $t = (T - T_c)/T_c$, and the external field $h = H/J$. Under coarse-graining, changing lengths to $x' = (1 - \epsilon)\, x$, the system is observed to be similar to itself at a shifted temperature $t' = (1 + a\epsilon)\, t$ and a shifted external field $h' = (1 + b\epsilon)\, h$, with ϵ infinitesimal and $a > b > 0$ (so there are two relevant eigendirections, with the temperature more strongly relevant than the external field). The curves shown below connect points that are similar up to some rescaling factor.
(a) *Which diagram below has curves consistent with this flow, for $a > b > 0$? Is the flow under coarse graining inward or outward from the origin?* (No math should be required. Hint: After coarse-graining, how does h/t change?)

The solid dots are at temperature t_0; the open circles are at temperature $t = 2t_0$.

(b) *In terms of ϵ and a, by what factor must x be rescaled by to relate the systems at t_0 and $2t_0$?* (Algebraic tricks: Use $(1+\delta) \approx \exp(\delta)$ everywhere. If you rescale multiple times until $\exp(na\epsilon) = 2$, you can solve for $(1-\epsilon)^n \approx \exp(-n\epsilon)$ without solving for n.) *If one of the solid dots in the appropriate figure from part (a) is at (t_0, h_0), what is the field $\widehat{h}$ for the corresponding open circle, in terms of a, b, ϵ, and the original coordinates?* (You may use the relation between $\widehat{h}$ and h_0 to check your answer for part (a).)

The magnetization $M(t,h)$ is observed to rescale under this same coarse-graining operation to $M' = (1+c\epsilon)\,M$, so $M\left((1+a\epsilon)\,t, (1+b\epsilon)\,h\right) = (1+c\epsilon)\,M(t,h)$.

(c) *Suppose $M(t,h)$ is known at (t_0,h_0), one of the solid dots. Give a formula for $M(2t_0, \widehat{h})$ at the corresponding open circle, in terms of $M(t_0,h_0)$, the original coordinates, a, b, c, and ϵ.* (Hint: Again, rescale n times.) *Substitute your formula for $\widehat{h}$ into the formula, and solve for $M(t_0,h_0)$.*

You have now basically derived the key result of the renormalization group; the magnetization curve at t_0 can be found from the magnetization curve at $2t_0$. In Section 12.2, we shall coarse-grain not to $t = 2t_0$, but to $t = 1$. We shall see that the magnetization everywhere can be predicted from the magnetization where the invariant curve crosses $t = 1$.

(d) *Substitute $2 \to 1/t_0$ in your formula from part (c). Show that $M(t,h) = t^{\beta}\mathcal{M}(h/t^{\beta\delta})$ (the standard scaling form for the magnetization in the Ising model). What are β and δ in terms of a, b, and c? How is $\mathcal{M}$ related to $M(t,h)$ where the curve crosses $t = 1$?*

(12.2) **Singular corrections to scaling.** (Condensed matter) ③

The renormalization group says that the number of relevant directions at the fixed point in system space is the number of parameters we need to tune to see a critical point, and that the critical exponents depend on the eigenvalues of these relevant directions. Do the irrelevant directions matter?

Let the Ising model in zero field be described by flow equations

$$\mathrm{d}t_\ell/\mathrm{d}\ell = t_\ell/\nu, \qquad \mathrm{d}u_\ell/\mathrm{d}\ell = -y u_\ell \tag{12.14}$$

where t_ℓ describes the renormalization of the reduced temperature $t = (T_c - T)/T_c$ after a coarse-graining by a factor $b = \exp(\ell)$, and u and u_ℓ represent a slowly-decaying irrelevant perturbation under the renormalization group. In Fig. 12.8, one may view t as the expanding eigendirection running roughly horizontally, and u as the contracting, irrelevant coordinate running roughly vertically. Thus our model starts with a value u associated to the distance in system space between R_c and S^*.

(a) *What is the invariant combination $z = ut^{\omega}$ that stays constant under the renormalization group? What is ω in terms of the eigenvalues $-y$ and $1/\nu$?*

Properties near critical points have universal power law singularities, but the corrections to these power laws also have universal properties predicted by the renormalization group. These come in two types – *analytic* corrections to scaling and *singular* corrections to scaling.

Let us consider corrections to the susceptibility. In analogy with other systems we have studied, we would expect that the susceptibility

$$\chi(t,u) = t^{-\gamma}X(z) \tag{12.15}$$

with $X(z)$ a universal function of the invariant combination you found in part (a). As a function of t, $\chi(t,u)$ has singularities at small t. But we expect properties to be analytic as we vary u, since the irrelevant direction is not being tuned to a special value, so we expect that a Taylor series of $\chi(t,u)$ in powers of u should make sense. Since $z \propto u$, we thus expect that $X(z)$ will be an analytic function of z for small z.[21]

(b) *Show that for small t, your z from part (a) goes to zero. Taylor expand $X(z)$. What corrections do you predict for the susceptibility from the first and second-order terms in the series?*

These are the singular corrections to scaling due to the irrelevant perturbation u.

An Ising magnet on a sample holder is loaded into a magnetometer, and the susceptibility is measured[22] at zero external field as a function

[21] Had we used a scaling variable $Z = tu^{1/\omega}$, for example, we would not have expected the corresponding scaling function to be analytic in small Z.

[22] The accuracy of the quoted exponents is not experimentally realistic.

of reduced temperature $t = (T - T_c)/T_c$. It is found to be well approximated by

$$\chi(T) = At^{-1.24} + Bt^{-0.83} + Ct^{0.42} + D + Et + \ldots \quad (12.16)$$

You may ignore any errors due to the magnetometer.

(c) *The exponent $\omega \approx 0.407$ for the 3D Ising universality class, and $\gamma \approx 1.237$. Which terms are explained as singular corrections to scaling?*

(d) *Can you provide a physical interpretation for the terms in eqn 12.16 that are not explained by your theory? For example, how do we expect the susceptibility of the sample holder to depend on temperature?*

So far, we have relied on universality and rescaling to derive the universal power laws and scaling forms for properties near critical points. We can derive these in a mathematical way by including the flows of the predictions along with the flows of the control parameters under coarse-graining:

$$\begin{aligned} \mathrm{d}\chi_\ell/\mathrm{d}\ell &= -(\gamma/\nu)\chi_\ell, \\ \mathrm{d}t_\ell/\mathrm{d}\ell &= t_\ell/\nu, \\ \mathrm{d}u_\ell/\mathrm{d}\ell &= -y u_\ell \end{aligned} \quad (12.17)$$

How do we derive the universal scaling function $X(z)$ from these renormalization group flows? Consider the flows illustrated in Fig. 12.8, except now with a third dimension involving the prediction χ. Consider a point (t_0, u_0, χ_0) in the system space, and the invariant curve defined by $z = u_0 t_0^{\omega}$ (dashed lines). Our renormalization group allows us to calculate $\chi_\ell(t_\ell, u_\ell)$ along these curves – relating the behavior everywhere near the critical manifold (vertical swath flowing toward S^*) to the properties along the outgoing trajectories, which approach closer and closer to the unstable manifold (the horizontal swath flowing away from S^*.

For example, we can define the universal scaling function $X(z)$ (for positive time t) to be the χ_{ℓ^*} where the flow crosses $t_{\ell^*} = 1$.

(e) *Solve eqns 12.17 for u_ℓ and t_ℓ. Setting $t_{\ell^*} = 1$, what is u_{ℓ^*} in terms of your invariant combination z?*

So we label each invariant scaling curve by the value of the vertical position u_{ℓ^*} where it crosses $t_{\ell^*} = 1$.

(f) *Solve eqns 12.17 for $\chi_{\ell^*}(1, u_{\ell^*})$, in terms of z, t_0, and $\chi_0(t_0, u_0)$. Use your solution to solve for the physical behavior $\chi_0(t_0, u_0)$ in terms of t and $X(z)$. Express $X(z)$ in terms of $\chi_{\ell^*}(1, u_{\ell^*})$.*

Remember the critical manifold is co-dimension one (or two, if you include temperature and external field), and the unstable manifold is dimension one (or two) – so we get universal predictions for a huge variety of systems, by observing the outgoing trajectories near a narrow tube or surface emitted from the fixed point.

(12.3) Nonlinear flows, analytic corrections, and hyperscaling.[23] ③

We consider the effects of nonlinear terms in renormalization group flows. The Ising model in zero field has one relevant variable (the deviation t of the temperature from T_c). To calculate the specific heat, we shall also consider the flow of the free energy per spin f under the renormalization group. Instead of a discrete coarse-graining by a factor b, here we use a continuous coarse-graining measured by ℓ. One can think of one coarse-graining step by $b = (1 + \epsilon)$ as incrementing $\ell \to \ell + \epsilon$; equivalently, coarse-graining to ℓ changes length scales by $\exp(\ell)$. Consider the particular flow equations[24]

$$\mathrm{d}f_\ell/\mathrm{d}\ell = Df_\ell - at_\ell{}^2 \qquad \mathrm{d}t_\ell/\mathrm{d}\ell = t_\ell/\nu, \quad (12.18)$$

where D is the dimension of space and $at_\ell{}^2$ is a nonlinear term that will be important in two dimensions.

Notice there are several free energies here. We shall call the free energy per spin of the actual system either f or f_0, and the temperature of the actual system either t or t_0. The coarse-grained free energy and temperature are f_ℓ and t_ℓ, after being coarse-grained by a factor $\exp(\ell)$. (Hence at $\ell = 0$ we have not yet coarse-grained, so $f_0 = f$ and $t_0 = t$.) Notice here that the free energy of our system is the *initial condition* $f_0(t_0)$ of this differential equation, and $f_\ell = f(\ell)$ and $t_\ell = t(\ell)$ are the renormalization group flows of the two variables.

[23] This exercise developed in collaboration with Colin Clement

[24] Note that these are total derivatives. So the first equation tells us the total change in f after coarsening by a factor $1 + \mathrm{d}\ell$. $f(t)$ then will coarse grain to $f_\ell(t_\ell)$ without needing to worry about the chain rule $\mathrm{d}f(t)/\mathrm{d}\ell = \partial f/\partial\ell + \partial f/\partial t \mathrm{d}t/\mathrm{d}\ell$.

To derive the scaling behavior, we shall coarse-grain to ℓ^* where $t_{\ell^*} = 1$, at which point the coarse-grained free energy is f_{ℓ^*}.

Let us start with the linear case $a = 0$.

(a) *Solve for f_ℓ and t_ℓ for $a = 0$. Setting $t_{\ell^*} = 1$, solve for f_0 in terms of f_{ℓ^*}, t_0, D, and ν. Solve for the specific heat per spin $c = T\partial^2 f/\partial T^2$, where $t = t_0 = (T - T_c)/T_c$ and $f = f_0$.* (Hint: Use the chain rule $\partial f/\partial T = (\partial f/\partial t)(\mathrm{d}t/\mathrm{d}T)$.) *Show that the specific heat near T_c has a power-law singularity $c \propto t^{-\alpha}$, with $\alpha = 2 - D\nu$.* (For example, in $D = 3$, $\nu \approx 0.63$, so $\alpha = 2 - D\nu \sim 0.11$; the specific heat diverges at T_c.) *Writing*

$$c = t^{-\alpha}(c_0 + c_1 t + c_2 t^2 + \ldots), \qquad (12.19)$$

what is the first correction c_1 to the specific heat near $t = 0$ in the absence of the nonlinear term?

Why is the linear term in the the free energy flow equal to the dimension, $\mathrm{d}f/\mathrm{d}\ell = Df + \ldots$, where all other terms are hard-to-compute critical exponents? There is no simple answer to this question.[25] Indeed, other models of disordered systems and glasses, and models above the upper critical dimension, the linear term is not given by the dimension. The relation $2 - \alpha = D\nu$ is called a *hyperscaling* relation (to emphasize they involve the dimension D), and these other models are said to violate hyperscaling.

(b) *In the case $a = 0$, show that the singular free energy f contained in a correlated volume ξ^D near the critical temperature becomes independent of the distance to the critical point.* (Hint: Look up the critical exponent describing how the correlation length ξ diverges as $t \to 0$.)

Glassy and disordered systems become extremely sluggish as they are cooled. In at least some cases, this is precisely because the energy barriers needed to continue equilibration diverge as their correlation lengths grow – they are glassy because their RG flows violate the hyperscaling relation.

So much for the power law singularity – what about the correction term c_1 in part (a)? It is an *analytic correction to scaling*.[26] Here it is *subdominant* – near the critical temperature where $t \to 0$, it is less singular than the leading term. One expects in a real physical system that the microscopic bond free energy J between spins will be some analytic function of temperature, and the physical free energy and specific heat will be multiplied by J. Expanding J in a Taylor series about $t = 0$ would also give us terms like those in eqn 12.19.

Does the introduction of the higher-order nonlinear term $at_\ell{}^2$ in eqn 12.18 change the behavior in an important way? Rather than exercising your expertise in analytic solutions of nonlinear differential equations, eqn 12.20 provides not f_ℓ and t_ℓ as functions of ℓ, but the relation between the two:

$$f_\ell(t_\ell) = f_0 \left(\frac{t_\ell}{t_0}\right)^{D\nu} - \frac{a\nu t_\ell{}^2}{(2 - D\nu)}(1 - (t_\ell/t_0)^{-(2-D\nu)}). \qquad (12.20)$$

(c) *Show that $f_\ell(t_\ell)$ in eqn 12.20 satisfies the differential equation given by eqn 12.18, using*

$$\frac{\mathrm{d}f_\ell}{\mathrm{d}t_\ell} = \frac{\mathrm{d}f_\ell}{\mathrm{d}\ell} \Big/ \frac{\mathrm{d}t_\ell}{\mathrm{d}\ell} \qquad (12.21)$$

Show that it has the correct initial conditions at $\ell = 0$. What is f_{ℓ^*} at ℓ^*, where $t_\ell = 1$? Show your method.

Again, it is important to remember that $f_\ell(t_\ell)$ is not the free energy as a function of temperature – it is the coarse-grained free energy as a function of the coarse-grained temperature of a system starting at a free energy f_0 at a temperature t_0. It is $f_0(t_0)$ that we want to know. Since here we have only one relevant variable (in zero field), all the flows lead to the same[27] final point $f_{\ell^*}(t_{\ell^*} = 1)$

[25] The total free energy of a system stays the same under coarse-graining (tracing out some degrees of freedom), and one can see that the free energy per spin thus must increase as $\mathrm{d}f_{\mathrm{coarsened}}/\mathrm{d}\ell = Df$ due to the reduction in the number of spins. But the RG has both coarse-graining and rescaling. Why is the total free energy not rescaled? Also, part of the coarse-grained free energy is left in an analytic background – why is the singular part governed by eqn 12.18 scaled to preserve the total singular free energy?

[26] These are distinct from *singular* corrections to scaling that arise, for example, from irrelevant terms under the renormalization group, that would produce subdominant corrections to c with powers $t^{-\alpha+\Delta}$ where Δ is not an integer, and is bigger than zero (hence subdominant).

[27] Remember for systems with more than one variable (say t and h), the free energy depends on the invariant curve departing from the fixed point (labeled, say, by $h/t^{\beta\delta} = h_{\ell^*}(t_{\ell^*} = 1)$. We solve for $f_0(t_0, h_0)$ in terms of $f_{\ell^*}(1, h_{\ell^*})$ just as we do in part (b).

(d) *Solve for f_0 in terms of f_{ℓ^*} and t_0. Solve for the specific heat $c = T\partial^2 f/\partial T^2$, where $t = (T - T_c)/T_c$. Show that it can be written in the form*

$$c = c_{+\text{analytic}}(t) + t^{-\alpha} c_{*\text{analytic}}(t) \qquad (12.22)$$

where the additive correction $c_{+\text{analytic}}(t)$ and the multiplicative correction $c_{*\text{analytic}}(t)$ have a simple Taylor series about $t = 0$. Write these two corrections, in terms of f_{ℓ^*}, T_c, a, ν, and D.

Here the nonlinear term a gives us not only a smooth multiplicative term in the specific heat, but also a smooth additive background. This clearly is also expected in a real physical system – other degrees of freedom uncoupled to the magnetization, or even the box holding the magnet, will contribute a specific heat that is analytic near T_c.

(12.4) **Beyond power laws: Nonlinear flows and logarithms in the 2D Ising model.**[28] ③

The two dimensional Ising model has a logarithmic singularity in the specific heat. The exact result shows that the specific heat per spin is

$$\begin{aligned} c(T) =& k_B \frac{2}{\pi}\left(\frac{2J}{k_B T_c}\right)^2 \big[-\log(1 - T/T_c) \\ &+ \log(k_B T_c/(2J) - (1 + \pi/4)\big] \\ =& -\frac{8}{\pi k_B T_c^2} \log\left(\frac{t}{\frac{1}{2} k_B T_c \exp(-(1+\pi/4))}\right) \\ =& -c_0 \log\left(\frac{t}{\tau}\right) \end{aligned} \qquad (12.23)$$

where $t = (T - T_c)/T_c$ and we set $J = 1$. (Remember $\log(t)$ is negative for small t.) Linearized flows around the renormalization group fixed point predict $c \sim t^{-\alpha}$, and when $\alpha \to 0$ one often observes logarithmic corrections. But such corrections are not predicted by the linearized flows. The key nonlinear term is the term $a{t_\ell}^2$ of eqn 12.18 in Exercise 12.3.

(a) *Is the solution for $f_\ell(t_\ell)$ in eqn 12.20 useful in $D = 2$? Why or why not?* (Hint: The exponent $\nu = 1$ for the two-dimensional Ising model.)

Again, we provide the solution of the nonlinear RG eqns 12.18 for $D = 2$

$$f_\ell(t_\ell) = f_0 (t_\ell/t_0)^2 + a{t_\ell}^2 \log(t_\ell/t_0). \qquad (12.24)$$

(b) *Show that $f_\ell(t_\ell)$ in eqn 12.24 satisfies the differential equation given by eqn 12.18, with the correct initial conditions. Solve for f_0 in terms of f_{ℓ^*} and t_0, where $t_{\ell^*} = 1$. Solve for the specific heat $c = T\partial^2 f/\partial T^2$, where $t = t_0 = (T - T_c)/T_c$ and $f = f_0$. (Remember the chain rule: $\partial f/\partial T = (\partial f/\partial t)(\mathrm{d}t/\mathrm{d}T)$.) Does it agree asymptotically with the exact result in eqn 12.23? What are c_0 and τ, in terms of a and f_{ℓ^*}?*

(12.5) **The Gutenberg Richter law.** (Scaling) ③

Power laws often arise at continuous transitions in non-equilibrium extended systems, particularly when disorder is important. We don't yet have a complete understanding of earthquakes, but they seem clearly related to the transition between pinned and sliding faults as the tectonic plates slide over one another.

The size S of an earthquake (the energy radiated, shown in the upper axis of Figure 12.3b) is traditionally measured in terms of a 'magnitude' $M \propto \log S$ (lower axis). The Gutenberg-Richter law tells us that the number of earthquakes of magnitude $M \propto \log S$ goes down as their size S increases. Figure 12.3b shows that the number of avalanches of magnitude between M and $M + 1$ is proportional to S^{-B} with $B \approx 2/3$. However, it is traditional in the physics community to consider the probability density $P(S)$ of having an avalanche of size S. *If $P(S) \sim S^{-\tau}$, give a formula for τ in terms of the Gutenberg-Richter exponent B.* (Hint: The bins in the histogram have different ranges of size S. Use $P(M)\,\mathrm{d}M = P(S)\,\mathrm{d}S$.)

(12.6) **Period Doubling.** (Scaling) ⓟ

Most of you will be familiar with the *period doubling route to chaos*, and the bifurcation diagram shown below. (See also Section 12.3.3).

[28] This exercise developed in collaboration with Colin Clement

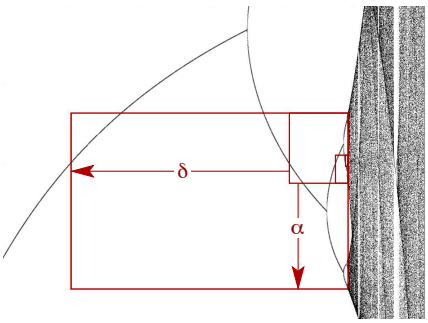


Fig. 12.20 Scaling in the period doubling bifurcation diagram. Shown are the points x on the attractor (vertical) as a function of the control parameter μ (horizontal), for the logistic map $f(x) = 4\mu x(1-x)$, near the transition to chaos.

The self-similarity here is not in space, but in time. It is discrete instead of continuous; the behavior is the similar if one rescales time by a factor of two, but not by a factor $1+\epsilon$. Hence instead of power laws we find a discrete self-similarity as we approach the critical point μ_∞.
(a) *From the diagram shown, roughly estimate the values of the Feigenbaum numbers δ (governing the rescaling of $\mu - \mu_\infty$) and α (govering the rescaling of $x - x_p$, where $x_p = 1/2$ is the peak of the logistic map).* (Hint: be sure to check the signs.)
Remember that the relaxation time for the Ising model became long near the critical temperature; it diverges as $t^{-\zeta}$ where t measures the distance to the critical temperature. Remember that the correlation length diverges as $t^{-\nu}$. Can we define ζ and ν for period doubling?
(b) *If each rescaling shown doubles the period T of the map, and T grows as $T \sim (\mu_\infty - \mu)^{-\zeta}$ near the onset of chaos, write ζ in terms of α and δ. If ξ is the* smallest *typical length scale of the attractor, and we define $\xi \sim (\mu_\infty - \mu)^{-\nu}$ (as is traditional at thermodynamic phase transitions), what is ν in terms of α and δ?* (Hint: be sure to check the signs.)

(12.7) **Random Walks.** (Scaling) ③
Self-similar behavior also emerges without proximity to any obvious transition. One might say that some *phases* naturally have self-similarity and power laws. Mathematicians have a technical term *generic* which roughly translates to 'without tuning a parameter to a special value', and so this is termed *generic scale invariance*.
The simplest example of generic scale invariance is that of a random walk. Figure 12.21 shows that a random walk appears statistically self-similar.

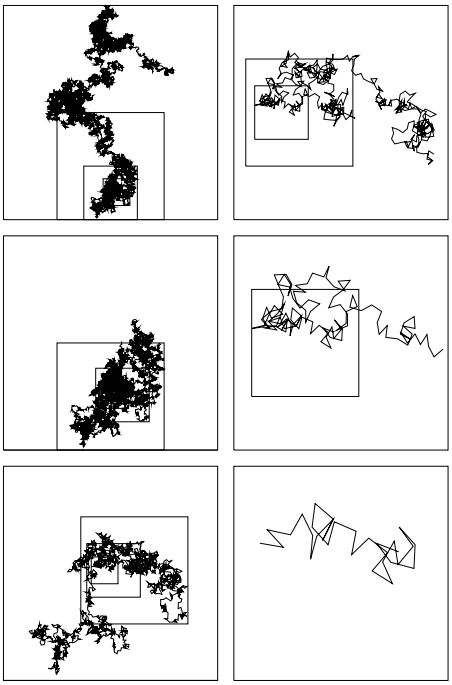


Fig. 12.21 Random walk scaling. Each box shows the first quarter of the random walk in the previous box. While each figure looks different in detail, they are *statistically* self-similar. That is, an ensemble of medium-length random walks would be indistinguishable from an ensemble of suitably rescaled long random walks.

Let $X(T) = \sum_{t=1}^{T} \xi_t$ be a random walk of length T, where ξ_t are independent random variables chosen from a distribution of mean zero and finite standard deviation. Derive the exponent ν governing the growth of the root-mean-square end-to-end distance $d(T) = \sqrt{\langle (X(T) - X(0))^2 \rangle}$ with T. Explain the connection between this and the formula from freshman lab courses for the way the standard deviation of the mean scales with the number of measurements.

(12.8) **Hysteresis and Barkhausen Noise.** (Scaling) ⓘ

Hysteresis is associated with abrupt phase transitions. Supercooling and superheating are examples (as temperature crosses T_c). Magnetic recording, the classic place where hysteresis is studied, is also governed by an abrupt phase transition – here the hysteresis in the magnetization, as the external field H is increased (to magnetize the system) and then decreased again to zero. Magnetic hysteresis is characterized by crackling (Barkhausen) electromagnetic noise. This noise is due to avalanches of spins flipping as the magnetic interfaces jerkily are pushed past defects by the external field (much like earthquake faults jerkily responding to the stresses from the tectonic plates). It is interesting that when dirt is added to this abrupt magnetic transition, it exhibits the power-law scaling characteristic of continuous transitions. Our model of magnetic hysteresis (unlike the experiments) has avalanches and scaling only at a special critical value of the disorder $R_c \sim 2.16$ (Figure 12.14). The integrated probability distribution $D(S,R)$ has a power law $D(S,R_c) \propto S^{-\bar{\tau}}$ at the critical point (where $\bar{\tau} = \tau + \sigma\beta\delta$ for our model) but away from the critical point takes the *scaling form*

$$D(S,R) \propto S^{-\bar{\tau}} \mathcal{D}(S^{\sigma}(R - R_c)). \tag{12.25}$$

Note from eqn (12.25) that at the critical disorder $R = R_c$ the distribution of avalanche sizes is a power law $D(S,R_c) = S^{-\bar{\tau}}$. The scaling form controls how this power law is altered as R moves away from the critical point. From Figure 12.14 we see that the main effect of moving above R_c is to cut off the largest avalanches at a typical largest size $S_{\max}(R)$, and another important effect is to form a 'bulge' of extra avalanches just below the cut–off.

Using the scaling form from eqn 12.25, with what exponent does $S_{\max}$ diverge as $r = (R_c - R) \to 0$? (Hint: At what size S is $D(S,R)$, say, one millionth of $S^{-\bar{\tau}}$?) *Given $\bar{\tau} \approx 2.03$, how does the mean $\langle S \rangle$ and the mean-square $\langle S^2 \rangle$ avalanche size scale with $r = (R_c - R)$?* (Hint: Your integral for the moments should have a lower cutoff S_0, the smallest possible avalanche, but no upper cutoff, since that is provided by the scaling function $\mathcal{D}$. Assume $\mathcal{D}(0) > 0$. Change variables to $Y = S^{\sigma} r$. Which moments diverge?)

(12.9) **First to fail: Weibull.**[29] (Mathematics, Statistics, Engineering) ③

Suppose you have a brand-new supercomputer with $N = 1000$ processors. Your parallelized code, which uses all the processors, cannot be restarted in mid-stream. How long a time t can you expect to run your code before the first processor fails?

This is example of *extreme value statistics* (see also exercises 12.10 and 12.11), where here we are looking for the smallest value of N random variables that are all bounded below by zero. For large N the probability distribution $\rho(t)$ and survival probability $S(t) = \int_t^{\infty} \rho(t')\,\mathrm{d}t'$ are often given by the *Weibull distribution*

$$\begin{aligned} S(t) &= \mathrm{e}^{-(t/\alpha)^{\gamma}}, \\ \rho(t) &= -\frac{\mathrm{d}S}{\mathrm{d}t} = \frac{\gamma}{\alpha}\left(\frac{t}{\alpha}\right)^{\gamma-1} \mathrm{e}^{-(t/\alpha)^{\gamma}}. \end{aligned} \tag{12.26}$$

Let us begin by assuming that the processors have a constant rate Γ of failure, so the probability density of a single processor failing at time t is $\rho_1(t) = \Gamma \exp(-\Gamma t)$ as $t \to 0$), and the survival probability for a single processor $S_1(t) = 1 - \int_0^t \rho_1(t')dt' \approx 1 - \Gamma t$ for short times.

(a) *Using $(1-\epsilon) \approx \exp(-\epsilon)$ for small ϵ, show that the the probability $S_N(t)$ at time t that all N processors are still running is of the Weibull form (eqn 12.26). What are α and γ?*

Often the probability of failure per unit time goes to zero or infinity at short times, rather than to a constant. Suppose the probability of failure for one of our processors

$$\rho_1(t) \sim Bt^k \tag{12.27}$$

with $k > -1$. (So, $k < 0$ might reflect a breaking-in period, where survival for the first few minutes increases the probability for later survival, and $k > 0$ would presume a dominant failure mechanism that gets worse as the processors wear out.)

(b) *Show the survival probability for N identical processors each with a power-law failure rate (eqn 12.27) is of the Weibull form for large N, and give α and γ as a function of B and k.*

The parameter α in the Weibull distribution just sets the scale or units for the variable t; only the exponent γ really changes the shape

[29] Developed with the assistance of Paul (Wash) Wawrzynek

of the distribution. Thus the form of the failure distribution at large N only depends upon the power law k for the failure of the individual components at short times, not on the behavior of $\rho_1(t)$ at longer times. This is a type of *universality*,[30] which here has a physical interpretation; at large N the system will break down soon, so only early times matter.

The Weibull distribution, we must mention, is often used in contexts not involving extremal statistics. Wind speeds, for example, are naturally always positive, and are conveniently fit by Weibull distributions.

Advanced discussion: Weibull and fracture toughness

Weibull developed his distribution when studying the fracture of materials under external stress. Instead of asking how long a time t a system will function, Weibull asked how big a load σ the material can support before it will snap.[31] Fracture in brittle materials often occurs due to pre-existing microcracks, typically on the surface of the material. Suppose we have an isolated[32] microcrack of length L in a (brittle) concrete pillar, lying perpendicular to the external stress. It will start to grow when the stress on the beam reaches a critical value roughly[33] given by

$$\sigma_c(L) \approx K_c/\sqrt{\pi L}. \tag{12.28}$$

Here K_c is the *critical stress intensity factor*, a material-dependent property which is high for steel and low for brittle materials like glass. (Cracks concentrate the externally applied stress σ at their tips into a square–root singularity; longer cracks have more stress to concentrate, leading to eqn 12.28.)

The failure stress for the material as a whole is given by the critical stress for the longest pre-existing microcrack. Suppose there are N microcracks in a beam. The length L of each microcrack has a probability distribution $\rho(L)$.

(c) *What is the probability distribution $\rho_1(\sigma)$ for the critical stress σ_c for a single microcrack, in terms of $\rho(L)$?* (Hint: Consider the population in a small range $\mathrm{d}\sigma$, and the same population in the corresponding range $\mathrm{d}\ell$.)

The distribution of microcrack lengths depends on how the material has been processed. The simplest choice, an exponential decay $\rho(L) \sim (1/L_0)\exp(-L/L_0)$, perversely does not yield a Weibull distribution, since the probability of a small critical stress does not vanish as a power law $B\sigma^k$ (eqn 12.27).

(d) *Show that an exponential decay of microcrack lengths leads to a probability distribution $\rho_1(\sigma)$ that decays faster than any power law at $\sigma = 0$ (i.e., is zero to all orders in σ).* (Hint: You may use the fact that e^x grows faster than x^m for any m as $x \to \infty$.)

Analyzing the distribution of failure stresses for a beam with N microcracks with this exponentially decaying length distribution yields a Gumbel distribution [138, section 8.2], not a Weibull distribution.

Many surface treatments, on the other hand, lead to power-law distributions of microcracks and other flaws, $\rho(L) \sim CL^{-\eta}$ with $\eta > 1$. (For example, fractal surfaces with power-law correlations arise naturally in models of corrosion, and on surfaces exposed by previous fractures.)

(e) *Given this form for the length distribution of microcracks, show that the distribution of fracture thresholds $\rho_1(\sigma) \propto \sigma^k$. What is k in terms of η?*

According to your calculation in part (b), this immediately implies a Weibull distribution of fracture strengths as the number of microcracks in the beam becomes large.

(12.10) **Biggest of bunch: Gumbel.** (Mathematics, Statistics, Engineering) ③

Much of statistical mechanics focuses on the average behavior in an ensemble, or the mean square fluctuations about that average. In

forms a one-parameter family of universality classes; see chapter 12.

[31] Many properties of a steel beam are largely independent of which beam is chosen. The elastic constants, the thermal conductivity, and the the specific heat depends to some or large extent on the morphology and defects in the steel, but nonetheless vary little from beam to beam—they are *self-averaging* properties, where the fluctuations due to the disorder average out for large systems. The fracture toughness of a given beam, however, will vary significantly from one steel beam to another. Self-averaging properties are dominated by the typical disordered regions in a material; fracture and failure are nucleated at the extreme point where the disorder makes the material weakest.

[32] The interactions between microcracks are often not small, and are a popular research topic.

[33] This formula assumes a homogeneous, isotropic medium as well as a crack orientation perpendicular to the external stress. In concrete, the microcracks will usually associated with grain boundaries, second-phase particles, porosity...

many cases, however, we are far more interested in the extremes of a distribution.

Engineers planning dike systems are interested in the highest flood level likely in the next hundred years. Let the high water mark in year j be H_j. Ignoring long-term weather changes (like global warming) and year-to-year correlations, let us assume that each H_j is an independent and identically distributed (IID) random variable with probability density $\rho_1(H_j)$. The *cumulative distribution function* (cdf) is the probability that a random variable is less than a given threshold. Let the cdf for a single year be $F_1(H) = P(H' < H) = \int^H \rho_1(H')\,\mathrm{d}H'$.

(a) *Write the probability $F_N(H)$ that the highest flood level (largest of the high-water marks) in the next $N = 1000$ years will be less than H, in terms of the probability $F_1(H)$ that the high-water mark in a single year is less than H.*

The distribution of the largest or smallest of N random numbers is described by *extreme value statistics* [138]. Extreme value statistics is a valuable tool in engineering (reliability, disaster preparation), in the insurance business, and recently in bioinformatics (where it is used to determine whether the best alignments of an unknown gene to known genes in other organisms are significantly better than that one would generate randomly).

(b) *Suppose that $\rho_1(H) = \exp(-H/H_0)/H_0$ decays as a simple exponential ($H > 0$). Using the formula*

$$(1 - A) \approx \exp(-A) \quad \text{small A} \tag{12.29}$$

show that the cumulative distribution function F_N for the highest flood after N years is

$$F_N(H) \approx \exp\left[-\exp\left(\frac{\mu - H}{\beta}\right)\right]. \tag{12.30}$$

for large H. (Why is the probability $F_N(H)$ small when H is not large, at large N?) What are μ and β for this case?

The constants β and μ just shift the scale and zero of the ruler used to measure the variable of interest. Thus, using a suitable ruler, the largest of many events is given by a Gumbel distribution

$$\begin{aligned} F(x) &= \exp(-\exp(-x)) \\ \rho(x) &= \partial F/\partial x = \exp(-(x + \exp(-x))). \end{aligned} \tag{12.31}$$

How much does the probability distribution for the largest of N IID random variables depend on the probability density of the individual random variables? Surprisingly little! It turns out that the largest of N Gaussian random variables also has the same Gumbel form that we found for exponentials. Indeed, any probability distribution that has unbounded possible values for the variable, but that decays faster than any power law, will have extreme value statistics governed by the Gumbel distribution [96, section 8.3]. In particular, suppose

$$F_1(H) \approx 1 - A\exp(-BH^\delta) \tag{12.32}$$

as $H \to \infty$ for some positive constants A, B, and δ. It is in the region near $H^*[N]$, defined by $F_1(H^*[N]) = 1 - 1/N$, that F_N varies in an interesting range (because of eqn 12.29).

(c) *Show that the extreme value statistics $F_N(H)$ for this distribution is of the Gumbel form (eqn 12.30) with $\mu = H^*[N]$ and $\beta = 1/(B\delta\, H^*[N]^{\delta-1})$.* (Hint: Taylor expand $F_1(H)$ at H^* to first order.)

The Gumbel distribution is *universal*. It describes the extreme values for any unbounded distribution whose tails decay faster than a power law.[34] (This is quite analogous to the central limit theorem, which shows that the normal or Gaussian distribution is the universal form for sums of large numbers of IID random variables, so long as the individual random variables have non-infinite variance.)

The Gaussian or standard normal distribution $\rho_1(H) = (1/\sqrt{2\pi})\exp(-H^2/2)$, for example, has a cumulative distribution $F_1(H) = (1/2)(1 + \mathrm{erf}(H/\sqrt{2}))$ which at large H has asymptotic form $F_1(H) \sim 1 - (1/\sqrt{2\pi}H)\exp(-H^2/2)$. This is of the general form of eqn 12.32 with $B = \frac{1}{2}$ and $\delta = 2$, except that A is a slowly varying function of H. This slow variation does not change the asymptotics. Hints for the numerics are available in the computer exercises section of the text Web site [131].

(d) *Generate $M = 10000$ lists of $N = 1000$ random numbers distributed with this Gaussian probability distribution. Plot a normalized histogram of the largest entries in each list. Plot also the predicted form $\rho_N(H) = \mathrm{d}F_N/\mathrm{d}H$*

[34] The Gumbel distribution can also describe extreme values for a bounded distribution, if the probability density at the boundary goes to zero faster than a power law [138, section 8.2].

from part (c). (Hint: $H^*(N) \approx 3.09023$ for $N = 1000$; check this if it is convenient.)
Other types of distributions can have extreme value statistics in different universality classes (see Exercise 12.11). Distributions with power-law tails (like the distributions of earthquakes and avalanches described in Chapter 12) have extreme value statistics described by *Fréchet distributions*. Distributions that have a strict upper or lower bound[35] have extreme value distributions that are described by Weibull statistics (see Exercise 12.9).

(12.11) **Extreme value statistics: Gumbel, Weibull, and Fréchet.** (Mathematics, Statistics, Engineering) ③
Extreme value statistics is the study of the maximum or minimum of a collection of random numbers. It has obvious applications in the insurance business (where one wants to know the biggest storm or flood in the next decades, see Exercise 12.10) and in the failure of large systems (where the weakest component or flaw leads to failure, see Exercise 12.9). Recently extreme value statistics has become of significant importance in bioinformatics. (In guessing the function of a new gene, one often searches entire genomes for good matches (or *alignments*) to the gene, presuming that the two genes are evolutionary descendents of a common ancestor and hence will have similar functions. One must understand extreme value statistics to evaluate whether the best matches are likely to arise simply at random.)
The limiting distribution of the biggest or smallest of N random numbers as $N \to \infty$ takes one of three *universal forms*, depending on the probability distribution of the individual random numbers. In this exercise we understand these forms as fixed points in a renormalization group.
Given a probability distribution $\rho_1(x)$, we define the *cumulative distribution function* (CDF) as $F_1(x) = \int_{-\infty}^{x} \rho(x')\,\mathrm{d}x'$. Let us define $\rho_N(x)$ to be the probability density that, out of N random variables, the largest is equal to x. Let $F_N(x)$ to be the corresponding CDF.
(a) *Write a formula for $F_{2N}(x)$ in terms of $F_N(x)$. If $F_N(x) = \exp(-g_N(x))$, show that $g_{2N}(x) = 2g_N(x)$.*
Our renormalization group coarse-graining operation will remove half of the variables, throwing away the smaller of every pair, and returning the resulting new probability distribution. In terms of the function $g(x) = -\log \int_{-\infty}^{x} \rho(x')dx'$, it therefore will return a rescaled version of the $2g(x)$. This rescaling is necessary because, as the sample size N increases, the maximum will drift upward—only the form of the probability distribution stays the same, the mean and width can change. Our renormalization-group coarse-graining operation thus maps function space into itself, and is of the form

$$T[g](x) = 2g(ax + b). \tag{12.33}$$

(This renormalization group is the same as that we use for sums of random variables in Exercise 12.24 where $g(k)$ is the logarithm of the Fourier transform of the probability density.)
There are three distinct types of fixed-point distributions for this renormalization group transformation, which (with an appropriate linear rescaling of the variable x) describe most extreme value statistics. The Gumbel distribution (Exercise 12.10) is of the form

$$\begin{aligned} F_{\mathrm{gumbel}}(x) &= \exp(-\exp(-x)) \\ \rho_{\mathrm{gumbel}}(x) &= \exp(-x)\exp(-\exp(-x)). \\ g_{\mathrm{gumbel}}(x) &= \exp(-x) \end{aligned}$$

The Weibull distribution (Exercise 12.9) is of the form

$$\begin{aligned} F_{\mathrm{weibull}}(x) &= \begin{cases} \exp(-(-x)^{\alpha}) & x < 0 \\ 1 & x \ge 0 \end{cases} \\ g_{\mathrm{weibull}}(x) &= \begin{cases} (-x)^{\alpha} & x < 0 \\ 0 & x \ge 0, \end{cases} \end{aligned} \tag{12.34}$$

and the Fréchet distribution is of the form

$$\begin{aligned} F_{\mathrm{fréchet}}(x) &= \begin{cases} 0 & x \le 0 \\ \exp(-x^{-\alpha}) & x > 0 \end{cases} \\ g_{\mathrm{fréchet}}(x) &= \begin{cases} \infty & x < 0 \\ x^{-\alpha} & x \ge 0, \end{cases} \end{aligned} \tag{12.35}$$

where $\alpha > 0$ in each case.
(b) *Show that these distributions are fixed points for our renormalization-group transformation eqn 12.33. What are a and b for each distribution, in terms of α?*

[35] ed distributions that have power-law asymptotics have Weibull statistics; see note 34 and Exer-

In parts (c) and (d) you will show that there are only these three fixed points $g^*(x)$ for the renormalization transformation, $T[g^*](x) = 2g^*(ax+b)$, up to an overall linear rescaling of the variable x, with some caveats...

(c) *First, let us consider the case $a \neq 1$. Show that the rescaling $x \to ax+b$ has a fixed point $x = \mu$. Show that the most general form for the fixed-point function is*

$$g^*(\mu \pm z) = z^{\alpha'} p_\pm(\gamma \log z) \tag{12.36}$$

for $z > 0$, where $p_\pm$ is periodic and α' and γ are constants such that $p_\pm$ has period equal to one. (Hint: Assume $p(y) \equiv 1$, find α', and then show $g^*/z^{\alpha'}$ is periodic.) *What are α' and γ? Which choice for a, p_+, and p_- gives the Weibull distribution? The Fréchet distribution?*

Normally the periodic function $p(\gamma \log(x-\mu))$ is assumed or found to be a constant (sometimes called $1/\beta$, or $1/\beta^{\alpha'}$). If it is not constant, then the probability density must have an infinite number of oscillations as $x \to \mu$, forming a weird essential singularity.

(d) *Now let us consider the case $a = 1$. Show again that the fixed-point function is*

$$g^*(x) = \mathrm{e}^{-x/\beta} p(x/\gamma) \tag{12.37}$$

with p periodic of period one, and with suitable constants β and γ. What are the constants in terms of b? What choice for p and β yields the Gumbel distribution?

Again, the periodic function p is often assumed a constant (e^{μ}), for reasons which are not as obvious as in part (c).

What are the domains of attraction of the three fixed points? If we want to study the maximum of many samples, and the initial probability distribution has $F(x)$ as its CDF, to which universal form will the extreme value statistics converge? Mathematicians have sorted out these questions. If $\rho(x)$ has a power-law tail, so $1 - F(x) \propto x^{-\alpha}$, then the extreme value statistics will be of the Frechet type, with the same α. If the initial probability distribution is bounded above at μ and if $1 - F(\mu - y) \propto y^{\alpha}$, then the extreme value statistics will be of the Weibull type. (More commonly, Weibull distributions arise as the smallest value from a distribution of positive random numbers, Exercise 12.9.) If the probability distribution decays faster than any polynomial (say, exponentially) then the extreme value statistics will be of the Gumbel form. (Gumbel extreme-value statistics can also arise for bounded random variables if the probability decays to zero faster than a power law at the bound.)

(12.12) **Diffusion equation and universal scaling functions.**[36] ②

The diffusion equation universally describes microscopic hopping systems at long length scales. We will investigate how to write the evolution in a universal scaling form.

The solution to a diffusion problem with a non-zero drift velocity is given by $\rho(x,t) = 1/\sqrt{4\pi Dt}\exp\left(-((x-vt)^2/(4Dt)\right)$. We will coarse grain by throwing away half the time points. We will then rescale the distribution so it looks like the original distribution. We can just write these two operations as $t' = t/2$, $x' = x/\sqrt{2}$, $\rho' = \sqrt{2}\rho$.[37] These three together constitute our renormalization group operation.

(a) *Write an expression for $\rho'(x',t')$ in terms of D, v, x', and t' (*not *in terms of D' and v'). Use it to determine the new renormalized velocity v' and diffusion constant D'. Are v and D relevant, irrelevant or marginal variables?*

Typically, whenever writing properties in a scaling function, there is some freedom in deciding which invariant combinations to use. Here let us use the invariant combination of variables, $\mathcal{X} = x/\sqrt{t}$ and $\mathcal{V} = \sqrt{t}v$. We can then write

$$\rho(x,t) = t^{-\alpha}\mathcal{P}(\mathcal{X},\mathcal{V},D), \tag{12.38}$$

a power law times a universal scaling function of invariant combination of variables.

(b) *Show that $\mathcal{X}$ and $\mathcal{V}$ are invariant under our renormalization group operation. What is α? Write an expression for $\mathcal{P}$, in terms of $\mathcal{X}$, $\mathcal{V}$, and D (and* not *x, v, or t).*

(Note that we need to solve the diffusion equation to find the universal scaling function $\mathcal{P}$, but we can learn a lot from just knowing that it is a fixed point of the renormalization group. So, the universal exponent α and the invariant scaling combinations $\mathcal{X}$, $\mathcal{V}$, and D are determined just by the coarsening and rescaling

[36] This exercise was developed in collaboration with Archishman Raju.
[37] Because ρ is a density, we need to rescale $\rho' \mathrm{d}x' = \rho \mathrm{d}x$.

steps in the renormalization group. In experiments and simulations, one often uses data to extract the universal critical exponents and universal scaling functions, relying on emergent scale invariance to tell us that a scaling form like eqn (12.38) is expected.)

(12.13) **Avalanche Size Distribution.** (Scaling function) ③

One can develop a mean-field theory for avalanches in non-equilibrium disordered systems by considering a system of N Ising spins coupled to one another by an infinite-range interaction of strength J/N, with an external field H and each spin also having a local random field h:

$$\mathcal{H} = -J_0/N \sum_{i,j} S_i S_j - \sum_i (H + h_i) S_i. \quad (12.39)$$

We assume that each spin flips over when it is pushed over; *i.e.*, when its change in energy

$$H_i^{\rm loc} = \frac{\partial \mathcal{H}}{\partial S_i} = J_0/N \sum_j S_j + H + h_i$$
$$= J_0 m + H + h_i$$

changes sign.[38] Here $m = (1/N)\sum_j S_j$ is the average magnetization of the system. All spins start by pointing down. A new avalanche is launched when the least stable spin (the unflipped spin of largest local field) is flipped by increasing the external field H. Each spin flip changes the magnetization by $2/N$. If the magnetization change from the first spin flip is enough to trigger the next-least-stable spin, the avalanche will continue.

We assume that the probability density for the random field $\rho(h)$ during our avalanche is a constant

$$\rho(h) = (1 + t)/(2J_0). \quad (12.40)$$

The constant t will measure how close the density is to the critical density $1/(2J_0)$.

(a) *Show that at $t = 0$ each spin flip will trigger on average one other spin to flip, for large N. Can you qualitatively explain the difference between the two phases with $t < 0$ and $t > 0$?*

We can solve exactly for the probability $D(S, t)$ of having an avalanche of size S. To have an avalanche of size S triggered by a spin with random field h, you must have precisely $S-1$ spins with random fields in the range $\{h, h+2J_0S/N\}$ (triggered when the magnetization changes by $2S/N$). The probability of this is given by the Poisson distribution. In addition, the random fields must be arranged so that the first spin triggers the rest. The probability of this turns out to be $1/S$.

(b) *(Optional) By imagining putting periodic boundary conditions on the interval $\{h, h + 2J_0S/N\}$, argue that exactly one spin out of the group of S spins will trigger the rest as a single avalanche.* (Hint from Ben Machta: For simplicity, we may assume[39] the avalanche starts at $H = m = 0$. Try plotting the local field $H^{\rm loc}(h') = J_0 m(h') + h'$ that a spin with random field h' would feel if the spins between h' and h were flipped. How would this function change if we shuffle all the random fields around the periodic boundary conditions?)

(c) *Show that the distribution of avalanche sizes is thus*

$$D(S, t) = \frac{S^{S-1}}{S!}(t + 1)^{S-1} \mathrm{e}^{-S(t+1)}. \quad (12.41)$$

With t small (near to the critical density) and for large avalanche sizes S we expect this to have a scaling form:

$$D(S, t) = S^{-\tau} \mathcal{D}(S/t^{-x}) \quad (12.42)$$

for some mean-field exponent x. That is, taking $t \to 0$ and $S \to \infty$ along a path with St^x fixed, we can expand $D(S, t)$ to find the scaling function.

(d) *Show that $\tau = 3/2$ and $x = 2$. What is the scaling function $\mathcal{D}$?* Hint: You'll need to use Stirling's formula $S! \sim \sqrt{2\pi S}(S/\mathrm{e})^S$ for large S, and that $1 + t = \exp(\log(1 + t)) \approx \mathrm{e}^{t - t^2/2 + t^3/3 \cdots}$.

This is a bit tricky to get right. Let's check it by doing the plots.

(e) *Plot $S^\tau D(S, t)$ versus $Y = S/t^{-x}$ for $t = 0.2$, 0.1, and 0.05 in the range $Y \in (0, 10)$. Does it converge to $\mathcal{D}(Y)$?*

(See *"Hysteresis and hierarchies, dynamics of disorder-driven first-order phase transformations"*, J. P. Sethna, K. Dahmen, S. Kartha,

[38] We ignore the self-interaction, which is unimportant at large N

[39] Equivalently, measure the random fields with respect to h_0 of the triggering spin, and let m be the magnetization change since the avalanche started.

J. A. Krumhansl, B. W. Roberts, and J. D. Shore, PRL **70**, 3347 (1993) for more information.)

(12.14) **Conformal Invariance.**[40] (Mathematics) ③ Emergence in physics describes new laws that arise from complex underpinnings, and often exhibits a larger symmetry than the original model. The diffusion equation emerges as a continuum limit from complex random microscopic motion, and diffusion on a square lattice has circular symmetry. Critical phenomena emerges near continuous phase transitions, and the resultant symmetry under dilations is exploited by the renormalization group to predict universal power laws and scaling functions.

The Ising model on a square lattice at the critical point, like the diffusion equation, also has an emergent circular symmetry: the complex patterns of up and down spins look the same on long length scales also when rotated by an angle. Indeed, making use of the symmetries under changes of length scale, position, and angle (plus one spatially nonuniform transformation), systems at their critical points have a *conformal* symmetry group.

In two dimensions, the conformal symmetry group becomes huge. Roughly speaking, any complex analytic function $f(z) = u(x+\mathrm{i}y) + \mathrm{i}v(x+\mathrm{i}y)$ takes a snapshot of an Ising model $M(x,y)$ and warps it into a new magnetization pattern at (u,v) that 'looks the same'. (Here u, v, x, and y are all real.)

You may remember that most ordinary functions (like z^2, $\sqrt{z}$, $\sin(z)$, $\log(z)$, and $\exp(\mathrm{i}z)$) are analytic. All of them yield cool transformations of the Ising model – weird and warped when magnified until you see the pixels, but recognizably Ising-like on long scales. This exercise will generate an example.

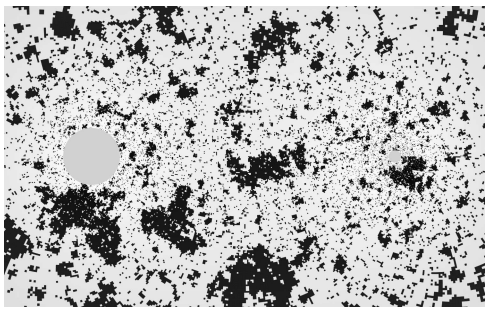


Fig. 12.22 Two proteins in a critical membrane. The figure shows the pixels of a critical Ising model simulated in a square, conformally warping the square onto the exterior of two circles (representing two proteins in a cell membrane). The warped pixels vary in size – largest in the upper and lower left, smallest near the smaller circle. They also locally rotate and translate the square lattice, but notice the pixels remain looking square – the angles and aspect ratios remain unchanged. The pixels are gray rather than black and white, with only the smallest pixels pure black and white; we must not only warp the pixels conformally, but also rescale the magnetization. Ignoring the pixelation, the different regions look statistically similar. A gray large pixel mimics the average color of similar-sized regions of tiny pixels.

(a) *What analytic function shrinks a region uniformly by a factor b, holding $z = 0$ fixed? What analytic function translates the lattice by a vector $\mathbf{r}_0 = (u_0, v_0)$? What analytic function rotates a region by an angle θ?*

(b) *Expanding $f(z+\delta) = f(z) + \delta f'(z)$, show that an analytic function f transforms a local region about z to linear order in δ by first rotating and dilating δ about z and then translating. What complex number gives the net translation?*

Figure 12.22 shows how one can use this conformal symmetry to study the interactions between circular 'proteins' embedded in a two dimensional membrane at an Ising critical point.[41]

In the renormalization group, we first coarse grain the system (shrinking by a factor b) and then rescale the magnetization (by some power b^{y_M}) in order to return to statistically the same critical state: $\widehat{M} = b^{y_M} M$. This rescaling turns

[40] This exercise is based on a project of Benjamin Machta's

[41] We use this transformation to study the effective interaction between two circular 'proteins' in a two-dimensional cell membrane near its critical point. The energy of attraction between two 'up' proteins is derivable from the energy of the square-lattice Ising model with the two side boundaries set to 'up'. (See B. B. Machta, S. L. Veatch, and J. P. Sethna, 'Critical Casimir forces in cellular membranes, *PRL* 109, 138101 (2012).)

the larger pixels in Fig. 12.22 more gray; a mostly up-spin 'white' region with tiny pixels is mimicked by a large single pixel with the statistically averaged gray color.

We can discover the correct power for M by examining the rescaling of the correlation function $C(\mathbf{r}) = \langle M(\mathbf{x})M(\mathbf{x}+\mathbf{r})\rangle$. In the Ising model at its critical point the correlation function $C(r) \sim r^{-(2-d+\eta)}$. In dimension $d = 2$, $\eta = 1/4$. We expect that the correlation function for the conformally transformed magnetization will be the same as the original correlation function.

(c) *If we coarse-grain by a constant factor b, what power of b must we multiply M by to make $C(\mathbf{r}) = \langle \widehat{M}(f(z))\widehat{M}(f(z)+\mathbf{r})\rangle$? Explain your reasoning.*

When our conformal transformation takes a pixel at $M(z)$ to a warped pixel of area A at $f(z)$, it rescales the magnetization by

$$\widehat{M}(f(z)) = |A|^{-1/16} M(z). \tag{12.43}$$

The pixel area for a locally uniform compression by b changes by $|\mathrm{d}f/\mathrm{d}z|^2 = 1/b^2$. You may use this to check your answer to part (c).

Figure 12.23 illustrates the self-similarity for the Ising model under rescaling, in analogy with Fig. 12.11's treatment of scale invariance in the random-field avalanche model. Here, unlike in Fig. 12.11, we incorporate the renormalization of the magnetization by changing the grayscale as we blow up successive lower–right–hand corners.

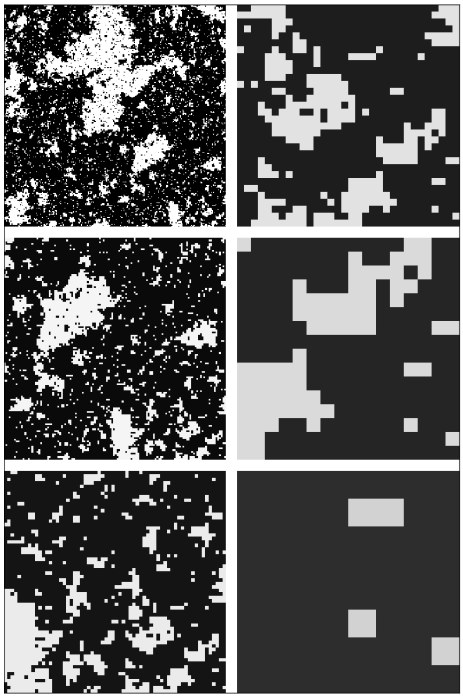


Fig. 12.23 Ising model under conformal rescaling. The 'powers of two' rescaling of the avalanches in Fig. 12.11 ignored this rescaling of M. Here we show the Ising model, again with the lowest right-hand quarter of each square inflated to fill the next – but now properly faded according to eqn 12.43.

Let us now explore the image of the Ising model under various analytic functions. Generate a snapshot of an Ising model, equilibrated at the critical temperature.[42]

Notice that Mathematica's notebook has trouble with plotting large numbers of polygons. Use Python. Or, if you prefer – save often, try not using a second display (a known problem), try not opening other programs in the background (including second Mathematica files), and avoid typing or otherwise disturbing the Mathematica window while it is running a long task. Debug everything with small systems, directly save your graphs of larger systems to a file, and display them outside Mathematica.

[42] You can get one from Matt Bierbaum's Web simulation at http://mattbierbaum.github.io/ising.js/) of size at least 256x256, or download one from the course Web site (along with hints files) at http://pages.physics.cornell.edu/~sethna/teaching/562/HW.html.

[43] If you work in Mathematica or Fortran, where the indices of arrays run from $(1 \ldots \mathrm{L})$, $z_{mn} = \left((m - \frac{1}{2}) + \mathrm{i}(n - \frac{1}{2})\right)/L$.

In either language, 512x512 will strain your machine; anything over 128x128 is acceptable, but larger systems will make the physics a bit clearer.

The hints files will allow you to import an Ising image, convert it into a two-dimensional array $S_{mn} = \pm 1$, and select an $L \times L$ subregion (if you wish, especially while you debug your code). We imagine these spins spread over the unit square in the complex plane; our code generates a list of spins and square polygons associated to each, with the spin S_{mn} sitting at the center[43] $z_{mn} = \left((m + \frac{1}{2}) + \mathrm{i}(n + \frac{1}{2})\right)/L$ of a $1/L \times 1/L$ square. The code will allow you to provide a function $f(z)$ of your choice, and will return the deformed quadrilaterals[44] centered at $f(z_{mn})$ with areas A_{mn}, and their associated rescaled spins $A_{mn}^{-1/16} S_{mn}$.[45] The software will also provide example routines showing how to output and shade in these quadrilaterals.

(d) *Generate a conformal transformation with a nonlinear analytic function of your choice, warping the Ising mode in an interesting way. Zoom in to regions occupied by lots of small pixels, where the individual pixels are not obvious.*[46] *Include both your entire plot and a cropped figure showing the expanded zoom. Discuss your zoomed plot critically – does it appear that the Ising correlations and visual structures have been faithfully retained by your analytic transformation?*

(e) *Load an Ising model equilibrated above the critical temperature at* $T = 100$ *(random noise), and one at* $T = 3$ *(short-range correlations). Distort and zoom each using your chosen conformal transformation. Analyze each and include your images. If you 'blur your eyes' enough to ignore the individual pixels, can you tell how much the system has been dilated? Are your conformally transformed images faithfully retaining the correlations and visual structures away from the critical point? For* $T = 3$*, which regions look qualitatively like* T_c*? Which regions look like* $T = 100$*?*

(f) *Invent a* non-analytic *function, and use it to distort your Ising model.* (Warning: most functions you write down, like $\log(\cosh^4(z + 1/z^2))$ will be analytic except at a few singularities. The author tried two methods: inventing real-valued functions $u(x, y)$ and $v(x, y)$ and forming $f = u + \mathrm{i}v$, and picking two different analytic functions $g(z)$ and $h(z)$ and using $f(z) = \mathrm{Re}(g(z)) + \mathrm{i}\mathrm{Im}(h(z))$. Make sure your function is non-analytic almost everywhere (e.g., violates the Cauchy-Riemann equations), not just at a point.)[47] *Find an example that makes for an interesting picture; include your images, including a zoom into a region with many pixels that range in size. As above, examine the images critically – does it appear that the Ising correlations and visual structures have been faithfully retained by your non-analytic transformation? Describe the distortions you see.* (Are the pixels still approximately square?)

Conformal symmetry in two dimensions was studied in an outgrowth of string theory. The representations of the conformal group allowed them to deduce the exact critical exponents for all the usual two dimensional statistical mechanical models – reproducing Onsager's result for the 2D Ising model, known solutions for the 2D tri-critical Ising model, 2D percolation, ... More recently, field theorists [71] have used conformal invariance in higher dimensions (with a strategy called 'conformal bootstrap') to produce bounds for critical exponents. They now hold the record on accuracy for the exponents of the 3D Ising model, giving $\beta = 0.326419(3)$, $\nu = 0.629971(4)$, and $\delta = 4.78984(1)$.

(12.15) **Ising self-similarity.** ⓘ

Start up the Ising model (computer exercises portion of the book web site [131]). Run a large

[44] The routine will drop quadrilaterals that extend to infinity, and also will remove quadrilaterals with 'n latter are associated with pixels which get inverted by $f(z)$; plotting packages usually do not provide rout exterior of a polygon.

[45] The list of quadrilaterals and spins will be linear, not two-dimensional, since graphics routines plotting pol flattened lists.

[46] You can zoom either using the graphics software or (sometimes more efficient) by saving a vector-graphics figure (like pdf) and viewing it separately. Start with $L = 64$ or so to make plots and zooming efficient, but for your final plots use L as large as feasible.

[47] Warning: The program automatically drops quadrilaterals with 'negative area', which usually happen when an internal point goes to infinity (and the polygon should be shaded 'outside'). This will also happen for the conjugate of an analytic function (e.g., $f(x + \mathrm{i}y) = y + \mathrm{i}x = \mathrm{i}\bar{z}$); you will get either errors about Transpose[] in Mathematica or an empty plot in Python. If this happens for your choice of $u(x, y) + \mathrm{i}v(x, y)$, try using $u - \mathrm{i}v$.

system at zero external field and $T = T_c = 2/\log(1+\sqrt{2}) \approx 2.26919$. Set the refresh rate low enough that graphics is not the bottle-neck, and run for at least a few hundred sweeps to equilibrate. You should see a fairly self-similar structure, with fractal-looking up-spin clusters inside larger down-spin structures inside ...
Can you find a nested chain of three clusters? Four?

(12.16) **Scaling and corrections to scaling.** (Condensed matter) ⓟ
Near critical points, the self-similarity under rescaling leads to characteristic power-law singularities. These dependences may be disguised, however, by less-singular corrections to scaling.
An experiment measures the susceptibility $\chi(T)$ in a magnet for temperatures T slightly above the ferromagnetic transition temperature T_c. They find their data is fit well by the form

$$\chi(T) = A(T-T_c)^{-1.25} + B + C(T-T_c) + D(T-T_c)^{1.77}. \tag{12.44}$$

(a) *Assuming this is the correct dependence near T_c, what is the critical exponent γ?*
When measuring functions of two variables near critical points, one finds universal scaling functions. The whole function is a prediction of the theory.
The pair correlation function $C(r,T) = \langle S(x)S(x+r)\rangle$ is measured in another, three-dimensional system just above T_c. It is found to be spherically symmetric, and of the form

$$C(r,T) = r^{-1.026}\mathcal{C}(r(T-T_c)^{0.65}), \tag{12.45}$$

where the function $\mathcal{C}(x)$ is found to be roughly $\exp(-x)$.
(b) *What is the critical exponent ν? The exponent η?*

(12.17) **Scaling and coarsening.** (Condensed matter) ⓟ
During coarsening, we found that the system changed with time, with a length scale that grows as a power of time: $L(t) \sim t^{1/2}$ for a non-conserved order parameter, and $L(t) \sim t^{1/3}$ for a conserved order parameter. These exponents, unlike critical exponents, are simple rational numbers that can be derived from arguments akin to dimensional analysis (Section 11.4.1). Associated with these diverging length scales there are scaling functions. Coarsening does not lead to a system which is self-similar to itself at equal times, but it does lead to a system which at two different times looks the same—apart from a shift of length scales.
An Ising model with non-conserved magnetization is quenched to a temperature T well below T_c. After a long time t_0, the correlation function looks like $C^{\mathrm{coar}}_{t_0}(\mathbf{r},T) = c(\mathbf{r})$.
Assume that the correlation function at short distances $C^{\mathrm{coar}}_t(\mathbf{0},T,t)$ will be time independent, and that the correlation function at later times will have the same functional form apart from a rescaling of the length. Write the correlation function at time twice t_0, $C^{\mathrm{coar}}_{2t_0}(\mathbf{r},T)$, in terms of $c(\mathbf{r})$. Write a scaling form

$$C^{\mathrm{coar}}_t(\mathbf{r},T) = t^{-\omega}\mathcal{C}(\mathbf{r}/t^{\rho},T). \tag{12.46}$$

Use the time independence of $C^{\mathrm{coar}}_t(\mathbf{0},T)$ and the fact that the order parameter is not conserved (Section 11.4.1) to predict the numerical values of the exponents ω and ρ.
It was only recently made clear that the scaling function $\mathcal{C}$ for coarsening *does* depend on temperature (and is, in particular, anisotropic for low temperature, with domain walls lining up with lattice planes). Low-temperature coarsening is not as 'universal' as continuous phase transitions are (Section 11.4.1); even in one model, different temperatures have different scaling functions.

(12.18) **Bifurcation theory.** (Mathematics) ⓘ
Dynamical systems theory is the study of the time evolution given by systems of differential equations. Let $\mathbf{x}(t)$ be a vector of variables evolving in time t, let $\boldsymbol{\lambda}$ be a vector of parameters governing the differential equation, and let $\mathbf{F}_\lambda(\mathbf{x})$ be the differential equations

$$\dot{\mathbf{x}} \equiv \frac{\partial \mathbf{x}}{\partial t} = \mathbf{F}_\lambda(\mathbf{x}). \tag{12.47}$$

The typical focus of the theory is not to solve the differential equations for general initial conditions, but to study the qualitative behavior. In general, they focus on *bifurcations*—special values of the parameters $\boldsymbol{\lambda}$ where the behavior of the system changes qualitatively.
(a) *Consider the differential equation in one variable $x(t)$ with one parameter μ:*

$$\dot{x} = \mu x - x^3. \tag{12.48}$$

Show that there is a bifurcation at $\mu_c = 0$, by showing that an initial condition with small,

non-zero $x(0)$ will evolve qualitatively differently at late times for $\mu > 0$ versus for $\mu < 0$. Hint: Although you can solve this differential equation explicitly, we recommend instead that you argue this qualitatively from the bifurcation diagram in Fig. 12.24; a few words should suffice.

Dynamical systems theory has much in common with equilibrium statistical mechanics of phases and phase transitions. The liquid–gas transition is characterized by external parameters $\boldsymbol{\lambda} = (P, T, N)$, and has a current state described by $\mathbf{x} = (V, E, \mu)$. Equilibrium phases correspond to fixed-points ($x^*(\mu)$ with $\dot{x}^* = 0$) in the dynamics, and phase transitions correspond to bifurcations.[48] For example, the power laws we find near continuous phase transitions have simpler analogues in the dynamical systems.

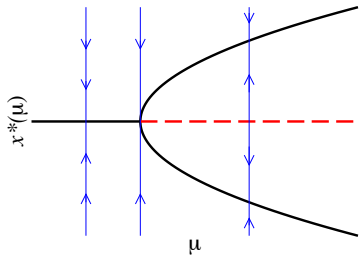


Fig. 12.24 Pitchfork bifurcation diagram. The flow diagram for the pitchfork bifurcation (eqn 12.48). The dashed line represents unstable fixed-points, and the solid thick lines represent stable fixed-points. The thin lines and arrows represent the dynamical evolution directions. It is called a pitchfork because of the three tines on the right emerging from the handle on the left.

(b) *Find the critical exponent β for the pitchfork bifurcation, defined by $x^*(\mu) \propto (\mu - \mu_c)^\beta$ as $\mu \to \mu_c$.*

Bifurcation theory also predicts universal behavior; all pitchfork bifurcations have the same scaling behavior near the transition.

(c) *At what value λ_c does the differential equation*

$$\dot{m} = \tanh(\lambda m) - m \tag{12.49}$$

have a bifurcation? Does the fixed-point value $m^*(\lambda)$ behave as a power law $m^* \sim |\lambda - \lambda_c|^\beta$ near λ_c (up to corrections with higher powers of $\lambda - \lambda_c$)? Does the value of β agree with that of the pitchfork bifurcation in eqn 12.48?

Just as there are different universality classes for continuous phase transitions with different renormalization-group fixed points, there are different classes of bifurcations each with its own *normal form*. Some of the other important normal forms include the saddle-node bifurcation,

$$\dot{x} = \mu - x^2, \tag{12.50}$$

transcritical exchange of stability,

$$\dot{x} = \mu x - x^2, \tag{12.51}$$

and the Hopf bifurcation,

$$\begin{aligned} \dot{x} &= (\mu - (x^2 + y^2))x - y, \\ \dot{y} &= (\mu - (x^2 + y^2))y + x. \end{aligned} \tag{12.52}$$

(12.19) **Mean-field theory.** (Condensed matter) ⓘ In Chapter 11 and Exercise 9.10, we make reference to mean-field theories, a term which is often loosely used for any theory which absorbs the fluctuations of the order parameter field into a single degree of freedom in an effective free energy. The original mean-field theory actually used the mean value of the field on neighboring sites to approximate their effects.

In the Ising model on a square lattice, this amounts to assuming each spin $s_j = \pm 1$ has four neighbors which are magnetized with the average magnetization $m = \langle s_j \rangle$, leading to a one-spin mean-field Hamiltonian

$$\mathcal{H} = -4Jms_j. \tag{12.53}$$

(a) *At temperature $k_B T$, what is the value for $\langle s_j \rangle$ in eqn 12.53, given m? At what temperature T_c is the phase transition, in mean field theory?* (Hint: At what temperature is a non-zero $m = \langle s \rangle$ self-consistent?) *Argue as in Exercise 12.18 part (c) that $m \propto (T_c - T)^\beta$ near T_c. Is this value for the critical exponent β correct for the Ising model in either two dimensions ($\beta = 1/8$) or three dimensions ($\beta \approx 0.325$)?*

(b) *Show that the mean-field solution you found in part (a) is the minimum in an effective*

[48] In Section 8.3, we noted that inside a phase all properties are analytic in the parameters. Similarly, bifurcations are values of λ where non-analyticities in the long-time dynamics are observed.

temperature-dependent free energy

$$V(m) = k_B T\Big(\frac{m^2}{2} - \log\left(\cosh(4Jm/k_BT)\right)\frac{k_BT}{4J}\Big). \tag{12.54}$$

On a single graph, plot $V(m)$ *for* $1/(k_BT) = 0.1$, 0.25, *and* 0.5, *for* $-2 < m < 2$, *showing the continuous phase transition.* Compare with Fig. 9.27.

(c) *What would the mean-field Hamiltonian be for the square-lattice Ising model in an external field* H*? Show that the mean-field magnetization is given by the minima in*[49]

$$V(m) = k_B T\Big(\frac{m^2}{2} - \log\left(\cosh((H+4Jm)/k_BT)\right)\frac{k_BT}{4J}\Big). \tag{12.55}$$

On a single graph, plot $V(m, H)$ *for* $\beta = 0.5$ *and* $H = 0$, 0.5, 1.0, *and* 1.5, *showing metastability and an abrupt transition. At what value of* H *does the metastable state become completely unstable?* Compare with Fig. 11.2(a).

(12.20) **The onset of lasing.**[50] (Quantum, Optics, Mathematics) ③

Lasers represent a stationary, condensed state. It is different from a phase of matter not only because it is made up out of energy, but also because it is intrinsically a non-equilibrium state. In a laser entropy is not maximized, free energies are not minimized—and yet the state has a robustness and integrity reminiscent of phases in equilibrium systems.

In this exercise, we will study a system of excited atoms coupled to a photon mode just before it begins to lase. We will see that it exhibits the diverging fluctuations and scaling that we have studied near critical points.

Let us consider a system of atoms weakly coupled to a photon mode. We assume that N_1 atoms are in a state with energy E_1, N_2 atoms are in a higher energy E_2, and that these atoms are strongly coupled to some environment that keeps these populations fixed.[51] Below the onset of lasing, the probability $\rho_n(t)$ that the photon mode is occupied by n photons obeys

$$\frac{\mathrm{d}\rho_n}{\mathrm{d}t} = a\big(n\rho_{n-1}N_2 - n\rho_nN_1 - (n+1)\rho_nN_2 + (n+1)\rho_{n+1}N_1\big). \tag{12.56}$$

The first term on the right-hand side represents the rate at which one of the N_2 excited atoms experiencing $n-1$ photons will emit a photon; the second term represents the rate at which one of the N_1 lower-energy atoms will absorb one of n photons; the third term represents emission in an environment with n photons, and the last represents absorption with $n+1$ photons. The fact that absorption in the presence of m photons is proportional to m and emission is proportional to $m+1$ is a property of bosons (Exercises 7.16(c) and 7.3). The constant $a > 0$ depends on the lifetime of the transition, and is related to the Einstein A coefficient (Exercise 7.16).

(a) *Find a simple expression for* $\mathrm{d}\langle n\rangle/\mathrm{d}t$, *where* $\langle n\rangle = \sum_{m=0}^{\infty} m\rho_m$ *is the mean number of photons in the mode.* (Hint: Collect all terms involving ρ_m.) *Show for* $N_2 > N_1$ *that this mean number grows indefinitely with time, leading to a macroscopic occupation of photons into this single state—a laser.*[52]

Now, let us consider our system just before it begins to lase. Let $\epsilon = (N_2 - N_1)/N_1$ be our measure of how close we are to the lasing instability. We might expect the value of $\langle n\rangle$ to diverge as $\epsilon \to 0$ like $\epsilon^{-\nu}$ for small ϵ. Near a phase transition, one also normally observes *critical slowing-down*: to equilibrate, the phase must communicate information over large distances of the order of the correlation length, which takes a time which diverges as the correlation length diverges. Let us define a critical-slowing-down exponent ζ for our lasing system, where the typical relaxation time is proportional to $|\epsilon|^{-\zeta}$ as $\epsilon \to 0$.

(b) *For* $\epsilon < 0$, *below the instability, solve your equation from part (a) for the long-time stationary value of* $\langle n\rangle$. *What is* ν *for our sys-*

[49] One must admit that it is a bit weird to have the external field H inside the effective potential, rather than coupled linearly to m outside.

[50] This exercise was developed with the help of Alex Gaeta and Al Sievers.

[51] That is, we assume that the atoms are being pumped into state N_2 to compensate for both decays into our photon mode and decays into other channels. This usually involves exciting atoms into additional atomic levels.

[52] The number of photons will eventually stop growing when they begin to pull energy out of the N_2 excited atoms faster than the pumping can replace them—invalidating our equations.

tem? For a general initial condition for the mean number of photons, solve for the time evolution. It should decay to the long-time value exponentially. Does the relaxation time diverge as $\epsilon \to 0$? What is ζ?

(c) *Solve for the stationary state $\boldsymbol{\rho}^*$ for $N_2 < N_1$.* (Your formula for ρ_n^* should not involve $\boldsymbol{\rho}^*$.) *If N_2/N_1 is given by a Boltzmann probability at temperature T, is $\boldsymbol{\rho}^*$ the thermal equilibrium distribution for the quantum harmonic oscillator at that temperature?* Warning: The number of bosons in a phonon mode is given by the Bose–Einstein distribution, but the probability of different occupations in a quantum harmonic oscillator is given by the Boltzmann distribution (see Section 7.2 and Exercise 7.11).

We might expect that near the instability the probability of getting n photons might have a scaling form

$$\rho_n^*(\epsilon) \sim n^{-\tau}\mathcal{D}(n|\epsilon|^{\nu}). \tag{12.57}$$

(d) *Show for small ϵ that there is a scaling form for $\boldsymbol{\rho}^*$, with corrections that go to zero as $\epsilon \to 0$, using your answer to part (c). What is τ? What is the function $\mathcal{D}(x)$?* (Hint: In deriving the form of $\mathcal{D}$, ϵ is small, but $n\epsilon^{\nu}$ is of order one. If you were an experimentalist doing scaling collapses, you would plot $n^{\tau}\rho_n$ versus $x = n|\epsilon|^{\nu}$; try changing variables in $n^{\tau}\rho_n$ to replace ϵ by x, and choose τ to eliminate n for small ϵ.)

(12.21) **Superconductivity and the renormalization group.** (Condensed matter) ⓘ

Ordinary superconductivity happens at a rather low temperature; in contrast to phonon energies (hundreds of degrees Kelvin times k_B) or electronic energies (tens of thousands of degrees Kelvin), phonon-mediated superconductivity in most materials happens below a few Kelvin. This is largely explained by the BCS theory of superconductivity, which predicts that the transition temperature for weakly-coupled superconductors is

$$T_c = 1.764\,\hbar\omega_D \exp\left(-1/Vg(\varepsilon_F)\right), \tag{12.58}$$

where ω_D is a characteristic phonon frequency, V is an attraction between electron pairs mediated by the phonons, and $g(\varepsilon_F)$ is the density of states (DOS) of the electron gas (eqn 7.74) at the Fermi energy. If V is small, $\exp\left(-1/Vg(\varepsilon_F)\right)$ can be exponentially small, explaining why materials often have to be so cold to go superconducting.

Superconductivity was discovered decades before it was explained. Many looked for explanations which would involve interactions with phonons, but there was a serious obstacle. People had studied the interactions of phonons with electrons, and had shown that the system stays metallic (no superconductivity) *to all orders in perturbation theory.*

(a) *Taylor expand T_c (eqn 12.58) about $V = 0^+$ (about infinitesimal positive V). Guess the value of all the terms in the Taylor series. Can we expect to explain superconductivity at positive temperatures by perturbing in powers of V?*

There are two messages here.

- Proving something to all orders in perturbation theory does not make it true.
- Since phases are regions in which perturbation theory converges (see Section 8.3), the theorem is not a surprise. It is a condition for a metallic phase with a Fermi surface to exist at all.

In recent times, people have developed a renormalization-group description of the Fermi liquid state and its instabilities[53] (see note 23 on p. 164). Discussing Fermi liquid theory, the BCS theory of superconductivity, or this renormalization-group description would take us far into rather technical subjects. However, we can illustrate all three by analyzing a rather unusual renormalization-group flow.

Roughly speaking, the renormalization-group treatment of Fermi liquids says that the Fermi surface is a fixed-point of a coarse-graining in *energy*. That is, they start with a system space consisting of a partially-filled band of electrons with an energy width W, including all kinds of possible electron–electron repulsions and attractions. They coarse-grain by perturbatively eliminating (integrating out) the electronic states near the edges of the band,

$$W' = (1-\delta)W, \tag{12.59}$$

incorporating their interactions and effects into altered interaction strengths among the remain-

[53] There are also other instabilities of Fermi liquids. Charge-density waves, for example, also have the characteristic $\exp(-1/aV)$ dependence on the coupling V.

ing electrons. These altered interactions give the renormalization-group flow in the system space. The equation for W gives the change under one iteration ($n = 1$); we can pretend n is a continuous variable and take $\delta n \to 0$, so $(W' - W)/\delta \to \mathrm{d}W/\mathrm{d}n$, and hence

$$\mathrm{d}W/\mathrm{d}n = -W. \tag{12.60}$$

When they do this calculation, they find the following.

- The non-interacting Fermi gas we studied in Section 7.7 is a *fixed point of the renormalization group*. All interactions are zero at this fixed-point. Let V represent one of these interactions.[54]
- The fixed-point is unstable to an attractive interaction $V > 0$, but is stable to a repulsive interaction $V < 0$.
- Attractive forces between electrons grow under coarse-graining and lead to new phases, but repulsive forces shrink under coarse-graining, leading back to the metallic free Fermi gas.

This is quite different from our renormalization-group treatment of phase transitions, where *relevant* directions like the temperature and field were unstable under coarse-graining, whether shifted up or down from the fixed-point, and other directions were *irrelevant* and stable (Fig. 12.8). For example, the temperature of our Fermi gas is a relevant variable, which rescales under coarse-graining like

$$\begin{aligned} T' &= (1 + a\delta)T, \\ \mathrm{d}T/\mathrm{d}n &= aT. \end{aligned} \tag{12.61}$$

Here $a > 0$, so the effective temperature becomes larger as the system is coarse-grained. How can they get a variable V which grows for $V > 0$ and shrinks for $V < 0$?

- When they do the coarse-graining, they find that the interaction V is *marginal*: to linear order it neither increases nor decreases.

The next allowed term in the Taylor series near the fixed-point gives us the coarse-grained equation for the interaction:

$$\begin{aligned} V' &= (1 + b\delta V)V, \\ \mathrm{d}V/\mathrm{d}n &= bV^2. \end{aligned} \tag{12.62}$$

- They find $b > 0$.

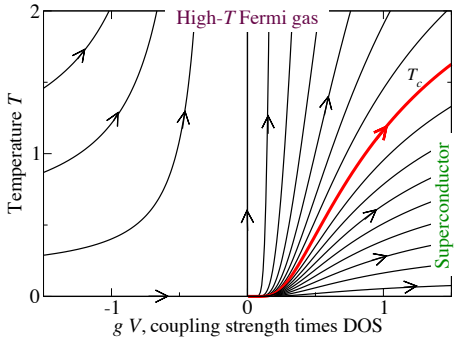


Fig. 12.25 Fermi liquid theory renormalization-group flows. The renormalization flows defined by eqns 12.61 and 12.62. The temperature T is relevant at the free Fermi gas fixed-point; the coupling V is marginal. The distinguished curve represents a phase transition boundary $T_c(V)$. Below T_c, for example, the system is superconducting; above T_c it is a (finite-temperature) metal.

(b) *True or false? (See Fig. 12.25.)*
(T) (F) For $V > 0$ (attractive interactions), the interactions get stronger with coarse-graining.
(T) (F) For $V < 0$ (repulsive interactions), coarse-graining leads us back to the free Fermi gas, explaining why the Fermi gas describes metals (Section 7.7).
(T) (F) Temperature is an irrelevant variable, but dangerous.
(T) (F) The scaling variable

$$x = TV^{1/\beta\delta} \tag{12.63}$$

is unchanged by the coarse-graining (second equations in 12.61 and 12.62), where β and δ are universal critical exponents;[55] *hence x labels the progress along the curves in Fig. 12.25 (increasing in the direction of the arrows).*
(T) (F) The scaling variable

$$y = T\exp\left(a/(bV)\right) \tag{12.64}$$

is unchanged by the coarse-graining, so each curve in Fig. 12.25 has a fixed value for y.

[54] V will be the pairing between opposite-spin electrons near the Fermi surface for superconductors.
[55] Note that here δ is not the infinitesimal change in parameter.

Now, without knowing anything about superconductivity, let us presume that our system goes superconducting at some temperature $T_c(V)$ when the interactions are attractive. When we coarse-grain a system that is at the superconducting transition temperature, we must get another system that is at its superconducting transition temperature.

(c) *What value for a/b must they calculate in order to get the BCS transition temperature (eqn 12.58) from this renormalization group? What is the value of the scaling variable (whichever you found in part (b)) along $T_c(V)$?*

Thus the form of the BCS transition temperature at small V, eqn 12.58, can be explained by studying the Fermi gas *without reference to the superconducting phase*!

(12.22) **Period doubling.**[56] (Mathematics, Complexity) ④

In this exercise, we use renormalization-group and scaling methods to study the *onset of chaos*. There are several routes by which a dynamical system can start exhibiting chaotic motion; this exercise studies the *period-doubling cascade*, first extensively investigated by Feigenbaum.

Chaos is often associated with dynamics which stretch and fold; when a batch of taffy is being pulled, the motion of a speck in the taffy depends sensitively on the initial conditions. A simple representation of this physics is provided by the map[57]

$$f(x) = 4\mu x(1-x) \tag{12.65}$$

restricted to the domain $(0,1)$. It takes $f(0) = f(1) = 0$, and $f(\frac{1}{2}) = \mu$. Thus, for $\mu = 1$ it precisely folds the unit interval in half, and stretches it to cover the original domain.

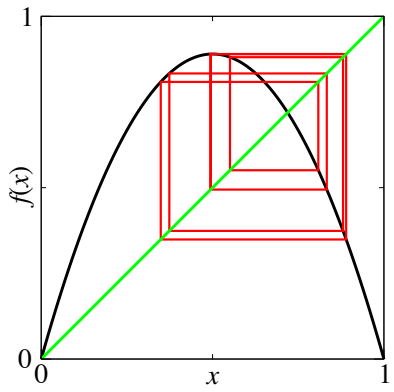


Fig. 12.26 Period-eight cycle. Iterating around the attractor of the Feigenbaum map at $\mu = 0.89$.

The study of dynamical systems (*e.g.*, differential equations and maps like eqn 12.65) often focuses on the behavior after long times, where the trajectory moves along the *attractor*. We can study the onset and behavior of chaos in our system by observing the evolution of the attractor as we change μ. For small enough μ, all points shrink to the origin; the origin is a stable fixed-point which attracts the entire interval $x \in (0,1)$. For larger μ, we first get a stable fixed-point inside the interval, and then *period doubling*.

(a) Iteration: *Set $\mu = 0.2$; iterate f for some initial points x_0 of your choosing, and convince yourself that they all are attracted to zero. Plot f and the diagonal $y = x$ on the same plot. Are there any fixed-points other than $x = 0$? Repeat for $\mu = 0.3$, $\mu = 0.7$, and 0.8. What happens? On the same graph, plot f, the diagonal $y = x$, and the segments $\{x_0, x_0\}$, $\{x_0, f(x_0)\}$, $\{f(x_0), f(x_0)\}$, $\{f(x_0), f(f(x_0))\}$, ... (representing the convergence of the trajectory to the attractor; see Fig. 12.26). See how $\mu = 0.7$ and 0.8 differ. Try other values of μ.*

By iterating the map many times, find a point a_0 on the attractor. As above, then plot the successive iterates of a_0 for $\mu = 0.7$, 0.8, 0.88, 0.89, 0.9, and 1.0.

You can see at higher μ that the system no longer settles into a stationary state at long times. The fixed-point where $f(x) = x$ ex-

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

[57] We also study this map in Exercises 4.10, 5.18, and 5.23; parts (a) and (b) below overlap somewhat with Exercise 4.10.

[58] In a continuous evolution, perturbations die away if the Jacobian of the derivative at the fixed-point has all negative eigenvalues. For mappings, perturbations die away if all eigenvalues of the Jacobian have magnitude less than one.

ists for all $\mu > 1/4$, but for larger μ it is no longer *stable*. If x^* is a fixed-point (so $f(x^*) = x^*$) we can add a small perturbation $f(x^* + \epsilon) \approx f(x^*) + f'(x^*)\epsilon = x^* + f'(x^*)\epsilon$; the fixed-point is stable (perturbations die away) if $|f'(x^*)| < 1$.[58]

In this particular case, once the fixed-point goes unstable the motion after many iterations becomes periodic, repeating itself after *two* iterations of the map—so $f(f(x))$ has two new fixed-points. This is called *period doubling*. Notice that by the chain rule $\mathrm{d}\,f(f(x))/\mathrm{d}x = f'(x)f'(f(x))$, and indeed

$$\begin{aligned}\frac{\mathrm{d}\,f^{[n]}}{\mathrm{d}x} &= \frac{\mathrm{d}\,f(f(\ldots f(x)\ldots))}{\mathrm{d}x} \\ &= f'(x)f'(f(x))\ldots f'(f(\ldots f(x)\ldots)),\end{aligned} \tag{12.66}$$

so the stability of a period-N orbit is determined by the product of the derivatives of f at each point along the orbit.

(b) Analytics: *Find the fixed-point $x^*(\mu)$ of the map 12.65, and show that it exists and is stable for $1/4 < \mu < 3/4$. If you are ambitious or have a computer algebra program, show that the period-two cycle is stable for $3/4 < \mu < (1+\sqrt{6})/4$.*

(c) Bifurcation diagram: *Plot the attractor as a function of μ, for $0 < \mu < 1$; compare with Fig. 12.17. (Pick regularly-spaced $\delta\mu$, run $n_{\mathrm{transient}}$ steps, record n_{cycles} steps, and plot. After the routine is working, you should be able to push $n_{\mathrm{transient}}$ and n_{cycles} both larger than 100, and $\delta\mu < 0.01$.) Also plot the attractor for another one-humped map*

$$f_{\sin}(x) = B\sin(\pi x), \tag{12.67}$$

for $0 < B < 1$. Do the bifurcation diagrams appear similar to one another?

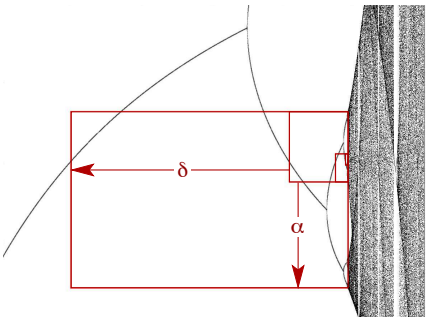


Fig. 12.27 Self-similarity in period-doubling bifurcations. The period doublings occur at geometrically-spaced values of the control parameter $\mu_\infty - \mu_n \propto \delta^{-n}$, and the attractor during the period-2^n cycle is similar to one-half of the attractor during the 2^{n+1}-cycle, except inverted and larger, rescaling x by a factor of α and μ by a factor of δ. The boxes shown in the diagram illustrate this self-similarity; each box looks like the next, except expanded by δ along the horizontal μ axis and flipped and expanded by α along the vertical axis.

Notice the complex, structured, chaotic region for large μ (which we study in Exercise 4.10). How do we get from a stable fixed-point $\mu < 3/4$ to chaos? The onset of chaos in this system occurs through a cascade of *period doublings*. There is the sequence of bifurcations as μ increases—the period-two cycle starting at $\mu_1 = 3/4$, followed by a period-four cycle starting at μ_2, period-eight at μ_3—a whole period-doubling cascade. The convergence appears geometrical, to a fixed-point μ_∞:

$$\mu_n \approx \mu_\infty - A\delta^{-n}, \tag{12.68}$$

so

$$\delta = \lim_{n\to\infty} (\mu_{n-1} - \mu_{n-2})/(\mu_n - \mu_{n-1}) \tag{12.69}$$

and there is a similar geometrical self-similarity along the x axis, with a (negative) scale factor α relating each generation of the tree (Fig. 12.27). In Exercise 4.10, we explained the boundaries in the chaotic region as images of $x = 1/2$. These special points are also convenient for studying period-doubling. Since $x = 1/2$ is the maximum in the curve, $f'(1/2) = 0$. If it were a fixed-point (as it is for $\mu = 1/2$), it would not only be stable, but unusually so: a shift by ϵ away from the fixed point converges after one step of the map to a distance $\epsilon f'(1/2) + \epsilon^2/2f''(1/2) = O(\epsilon^2)$. We say that such a fixed-point is *superstable*. If we have a period-N orbit that passes through $x = 1/2$, so that the Nth iterate $f^N(1/2) \equiv f(\ldots f(1/2)\ldots) = 1/2$, then the orbit is also superstable, since (by eqn 12.66) the derivative of the iterated map is the product of the derivatives along the orbit, and hence is also zero.

These superstable points happen roughly halfway between the period-doubling bifurcations, and are easier to locate, since we know that

$x = \frac{1}{2}$ is on the orbit. Let us use them to investigate the geometrical convergence and self-similarity of the period-doubling bifurcation diagram from part (d). For this part and part (h), you will need a routine that finds the roots $G(y) = 0$ for functions G of one variable y.

(d) The Feigenbaum numbers and universality: *Numerically, find the values of μ_n^s at which the 2^n-cycle is superstable, for the first few values of n.* (Hint: Define a function $G(\mu) = f_\mu^{[2^n]}(\frac{1}{2}) - \frac{1}{2}$, and find the root as a function of μ. In searching for μ_{n+1}^s, you will want to search in a range $(\mu_n^s + \epsilon, \mu_n^s + (\mu_n^s - \mu_{n-1}^s)/A)$ where $A \sim 3$ works pretty well. Calculate μ_0 and μ_1 by hand.) *Calculate the ratios $(\mu_{n-1}^s - \mu_{n-2}^s)/(\mu_n^s - \mu_{n-1}^s)$; do they appear to converge to the Feigenbaum number $\delta = 4.6692016091029909\ldots$? Extrapolate the series to μ_∞ by using your last two reliable values of μ_n^s and eqn 12.69. In the superstable orbit with 2^n points, the nearest point to $x = \frac{1}{2}$ is $f^{[2^{n-1}]}(\frac{1}{2})$.*[59] *Calculate the ratios of the amplitudes $f^{[2^{n-1}]}(\frac{1}{2}) - \frac{1}{2}$ at successive values of n; do they appear to converge to the universal value $\alpha = -2.50290787509589284\ldots$? Calculate the same ratios for the map $f_2(x) = B\sin(\pi x)$; do α and δ appear to be universal (independent of the mapping)?*

The limits α and δ are independent of the map, so long as it folds (one hump) with a quadratic maximum. They are the same, also, for experimental systems with many degrees of freedom which undergo the period-doubling cascade. This self-similarity and universality suggests that we should look for a renormalization-group explanation.

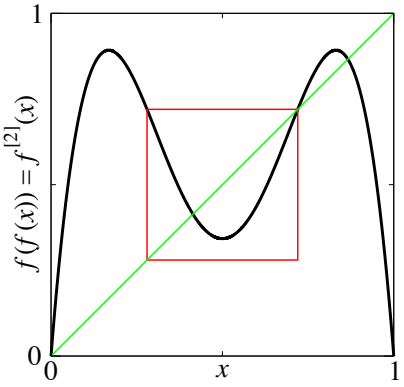


Fig. 12.28 Renormalization-group transformation. The renormalization-group transformation takes $g(g(x))$ in the small window with upper corner x^* and inverts and stretches it to fill the whole initial domain and range $(0,1) \times (0,1)$.

(e) Coarse-graining in time. *Plot $f(f(x))$ vs. x for $\mu = 0.8$, together with the line $y = x$ (or see Fig. 12.28). Notice that the period-two cycle of f becomes a pair of stable fixed-points for $f^{[2]}$.* (We are coarse-graining in time—removing every other point in the time series, by studying $f(f(x))$ rather than f.) *Compare the plot with that for $f(x)$ vs. x for $\mu = 0.5$. Notice that the region zoomed in around $x = \frac{1}{2}$ for $f^{[2]} = f(f(x))$ looks quite a bit like the entire map f at the smaller value $\mu = 0.5$. Plot $f^{[4]}(x)$ at $\mu = 0.875$; notice again the small one-humped map near $x = \frac{1}{2}$.*

The fact that the one-humped map reappears in smaller form just after the period-doubling bifurcation is the basic reason that succeeding bifurcations so often follow one another. The fact that many things are universal is due to the fact that the little one-humped maps have a shape which becomes *independent of the original map* after several period-doublings.

Let us define this renormalization-group transformation T, taking function space into itself. Roughly speaking, T will take the small upside-down hump in $f(f(x))$ (Fig. 12.28), invert it, and stretch it to cover the interval from $(0,1)$. Notice in your graphs for part (g) that the line $y = x$ crosses the plot $f(f(x))$ not only at the two points on the period-two attrac-

[59] This is true because, at the previous superstable orbit, 2^{n-1} iterates returned us to the original point $x = \frac{1}{2}$.

[60] For asymmetric maps, we would need to locate this other corner $f(f(x_c)) = x^*$ numerically. As it happens, breaking this symmetry is irrelevant at the fixed-point.

tor, but also (naturally) at the old fixed-point $x^*[f]$ for $f(x)$. This unstable fixed-point plays the role for $f^{[2]}$ that the origin played for f; our renormalization-group rescaling must map $(x^*[f], f(x^*)) = (x^*, x^*)$ to the origin. The corner of the window that maps to $(1,0)$ is conveniently located at $1-x^*$, since our map happens to be symmetric[60] about $x = \frac{1}{2}$. For a general one-humped map $g(x)$ with fixed-point $x^*[g]$ the side of the window is thus of length $2(x^*[g]-\frac{1}{2})$. To invert and stretch, we must thus rescale by a factor $\alpha[g] = -1/(2(x^*[g] - \frac{1}{2}))$. Our renormalization-group transformation is thus a mapping $T[g]$ taking function space into itself, where

$$T[g](x) = \alpha[g]\,(g\,(g(x/\alpha[g] + x^*[g])) - x^*[g])\,. \tag{12.70}$$

(This is just rescaling x to squeeze into the window, applying g twice, shifting the corner of the window to the origin, and then rescaling by α to fill the original range $(0,1)\times(0,1)$.)

(f) Scaling and the renormalization group: *Write routines that calculate $x^*[g]$ and $\alpha[g]$, and define the renormalization-group transformation $T[g]$. Plot $T[f]$, $T[T[f]]$,... and compare them. Are we approaching a fixed-point f^* in function space?*

This explains the self-similarity; in particular, the value of $\alpha[g]$ as g iterates to f^* becomes the Feigenbaum number $\alpha = -2.5029\ldots$.

(g) Universality and the renormalization group: *Using the sine function of eqn 12.67, compare $T[T[f_{\sin}]]$ to $T[T[f]]$ at their onsets of chaos. Are they approaching the same fixed-point?*

By using this rapid convergence in function space, one can prove both that there will (often) be an infinite geometrical series of period-doubling bifurcations leading to chaos, and that this series will share universal features (exponents α and δ and features) that are independent of the original dynamics.

(12.23) **The renormalization group and the central limit theorem: short.** (Mathematics) ④

If you are familiar with the renormalization group and Fourier transforms, this problem can be stated very quickly. If not, you are probably better off doing the long version (Exercise 12.24).

Write a renormalization-group transformation T taking the space of probability distributions into itself, that takes two random variables, adds them, and rescales the width by the square root of two [28]. Show that the Gaussian of width σ is a fixed-point. Find the eigenfunctions f_n and eigenvectors λ_n of the linearization of T at the fixed-point. (Hint: It is easier in Fourier space.) *Describe physically what the relevant and marginal eigenfunctions represent. By subtracting the fixed-point distribution from a binomial distribution, find the leading correction to scaling, as a function of x. Which eigenfunction does it represent? Why is the leading irrelevant eigenvalue not dominant here?*

(12.24) **The renormalization group and the central limit theorem: long.** (Mathematics) ④

In this exercise, we will develop a renormalization group in *function space* to derive the central limit theorem [28]. We will be using maps (like our renormalization transformation T) that take a function ρ of x into another function of x; we will write $T[\rho]$ as the new function, and $T[\rho](x)$ as the function evaluated at x. We will also make use of the Fourier transform (eqn A.6)

$$\mathcal{F}[\rho](k) = \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}kx}\rho(x)\,\mathrm{d}x; \tag{12.71}$$

$\mathcal{F}$ maps functions of x into functions of k. When convenient, we will also use the tilde notation: $\widetilde{\rho} = \mathcal{F}[\rho]$, so for example (eqn A.7)

$$\rho(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}kx}\widetilde{\rho}(k)\,\mathrm{d}k. \tag{12.72}$$

The central limit theorem states that the sum of many independent random variables tends to a Gaussian, whatever the original distribution might have looked like. That is, the Gaussian distribution is the *fixed-point function* for large sums. When summing many random numbers, the details of the distributions of the individual random variables becomes unimportant; simple behavior emerges. We will study this using the renormalization group, giving an example where we can explicitly implement the coarse-graining transformation. Here our system space is the space of probability distributions $\rho(x)$. There are four steps in the procedure.

1. *Coarse-grain.* Remove some fraction (usually half) of the degrees of freedom. Here, we will add pairs of random variables; the probability distribution for sums of N independent random

variables of distribution f is the same as the distribution for sums of $N/2$ random variables of distribution $f * f$, where $*$ denotes convolution.
(a) *Argue that if $\rho(x)$ is the probability that a random variable has value x, then the probability distribution of the sum of two random variables drawn from this distribution is the convolution*

$$C[\rho](x) = (\rho * \rho)(x) = \int_{-\infty}^{\infty} \rho(x-y)\rho(y)\,\mathrm{d}y. \tag{12.73}$$

Remember (eqn A.23) the Fourier transform of the convolution is the product of the Fourier transforms, so

$$\mathcal{F}[C[\rho]](k) = \left(\widetilde{\rho}(k)\right)^2. \tag{12.74}$$

2. *Rescale.* The behavior at larger lengths will typically be similar to that of smaller lengths, but some of the constants will shift (or *renormalize*). Here the mean and width of the distributions will increase as we coarse-grain. We confine our main attention to distributions of zero mean. Remember that the width (standard deviation) of the sum of two random variables drawn from ρ will be $\sqrt{2}$ times the width of one variable drawn from ρ, and that the overall height will have to shrink by $\sqrt{2}$ to stay normalized. We define a rescaling operator $S_{\sqrt{2}}$ which reverses this spreading of the probability distribution:

$$S_{\sqrt{2}}[\rho](x) = \sqrt{2}\rho(\sqrt{2}x). \tag{12.75}$$

(b) *Show that if ρ is normalized (integrates to one), so is $S_{\sqrt{2}}[\rho]$. Show that the Fourier transform is*

$$\mathcal{F}[S_{\sqrt{2}}[\rho]](k) = \widetilde{\rho}(k/\sqrt{2}). \tag{12.76}$$

Our renormalization-group transformation is the composition of these two operations,

$$\begin{aligned} T[\rho](x) &= S_{\sqrt{2}}[C[\rho]](x) \\ &= \sqrt{2}\int_{-\infty}^{\infty} \rho(\sqrt{2}x-y)\rho(y)\,\mathrm{d}y. \end{aligned} \tag{12.77}$$

Adding two Gaussian random variables (convolving their distributions) and rescaling the width back should give the original Gaussian distribution; the Gaussian should be a *fixed-point*.
(c) *Show that the Gaussian distribution*

$$\rho^*(x) = (1/\sqrt{2\pi}\sigma)\exp(-x^2/2\sigma^2) \tag{12.78}$$

is indeed a fixed-point in function space under the operation T. You can do this either by direct integration, or by using the known properties of the Gaussian under convolution.
(d) *Use eqns 12.74 and 12.76 to show that*

$$\mathcal{F}[T[\rho]](k) = \widetilde{T}[\widetilde{\rho}](k) = \widetilde{\rho}(k/\sqrt{2})^2. \tag{12.79}$$

Calculate the Fourier transform of the fixed-point $\widetilde{\rho^*}(k)$ (or see Exercise A.4). Using eqn 12.79, show that $\widetilde{\rho^*}(k)$ is a fixed-point in Fourier space under our coarse-graining operator $\widetilde{T}$.[61]
These properties of T and ρ^* should allow you to do most of the rest of the exercise without any messy integrals.
The central limit theorem tells us that sums of random variables have probability distributions that approach Gaussians. In our renormalization-group framework, to prove this we might try to show that our Gaussian fixed-point is *attracting*: that all nearby probability distributions flow under iterations of T to ρ^*.
3. *Linearize about the fixed point.* Consider a function near the fixed point: $\rho(x) = \rho^*(x) + \epsilon f(x)$. In Fourier space, $\widetilde{\rho}(k) = \widetilde{\rho^*}(k) + \epsilon\widetilde{f}(k)$. We want to find the eigenvalues λ_n and eigenfunctions f_n of the derivative of the mapping T. That is, they must satisfy

$$\begin{aligned} T[\rho^* + \epsilon f_n] &= \rho^* + \lambda_n \epsilon f_n + O(\epsilon^2), \\ \widetilde{T}[\widetilde{\rho^*} + \epsilon\widetilde{f_n}] &= \widetilde{\rho^*} + \lambda_n\epsilon\widetilde{f_n} + O(\epsilon^2). \end{aligned} \tag{12.80}$$

(e) *Show using eqns 12.79 and 12.80 that the transforms of the eigenfunctions satisfy*

$$\widetilde{f_n}(k) = (2/\lambda_n)\widetilde{\rho^*}(k/\sqrt{2})\widetilde{f_n}(k/\sqrt{2}). \tag{12.81}$$

4. *Find the eigenvalues and calculate the universal critical exponents.*
(f) *Show that*

$$\widetilde{f_n}(k) = (\mathrm{i}k)^n\widetilde{\rho^*}(k) \tag{12.82}$$

is the Fourier transform of an eigenfunction (i.e., that it satisfies eqn 12.81.) What is the eigenvalue λ_n?

[61] To be explicit, the operator $\widetilde{T} = \mathcal{F} \circ T \circ \mathcal{F}^{-1}$ is a renormalization-group transformation that maps Fourier space into itself.

Our fixed-point actually does not attract all distributions near it. The directions with eigenvalues greater than one are called *relevant*; they are dangerous, corresponding to deviations from our fixed-point that grow under coarse-graining. The directions with eigenvalues equal to one are called *marginal*; they do not get smaller (to linear order) and are thus also potentially dangerous. When you find relevant and marginal operators, you always need to understand each of them on physical grounds.

(g) *The eigenfunction $f_0(x)$ with the biggest eigenvalue corresponds to an unphysical perturbation; why?* (Hint: Probability distributions must be normalized to one.) *The next two eigenfunctions f_1 and f_2 have important physical interpretations. Show that $\rho^* + \epsilon f_1$ to lowest order is equivalent to a shift in the mean of ρ, and $\rho^* + \epsilon f_2$ is a shift in the standard deviation σ of ρ^*.*

In this case, the relevant perturbations do not take us to qualitatively new phases—just to other Gaussians with different means and variances. All other eigenfunctions should have eigenvalues λ_n less than one. This means that a perturbation in that direction will shrink under the renormalization-group transformation:

$$T^N(\rho^* + \epsilon f_n) - \rho^* \sim \lambda_n^N \epsilon f_n. \tag{12.83}$$

Corrections to scaling and coin flips. Does anything really new come from all this analysis? One nice thing that comes out is the *leading corrections to scaling*. The fixed-point of the renormalization group explains the Gaussian shape of the distribution of N coin flips in the limit $N \to \infty$, but the linearization about the fixed-point gives a systematic understanding of the corrections to the Gaussian distribution for large but not infinite N.

Usually, the largest eigenvalues are the ones which dominate. In our problem, consider adding a small perturbation to the fixed-point f^* along the two leading irrelevant directions f_3 and f_4:

$$\rho(x) = \rho^*(x) + \epsilon_3 f_3(x) + \epsilon_4 f_4(x). \tag{12.84}$$

These two eigenfunctions can be inverse-transformed from their k-space form (eqn 12.82):

$$\begin{aligned} f_3(x) &\propto \rho^*(x)(3x/\sigma - x^3/\sigma^3), \\ f_4(x) &\propto \rho^*(x)(3 - 6x^2/\sigma^2 + x^4/\sigma^4). \end{aligned} \tag{12.85}$$

What happens to these perturbations under multiple applications of our renormalization-group transformation T? After ℓ applications (corresponding to adding together 2^ℓ of our random variables), the new distribution should be given by

$$T^\ell(\rho)(x) \sim \rho^*(x) + \lambda_3^\ell \epsilon_3 f_3(x) + \lambda_4^\ell \epsilon_4 f_4(x). \tag{12.86}$$

Since $1 > \lambda_3 > \lambda_4 \ldots$, the leading correction should be dominated by the perturbation with the largest eigenvalue.

(h) *Plot the difference between the binomial distribution giving the probability of m heads in N coin flips, and a Gaussian of the same mean and width, for $N = 10$ and $N = 20$.* (The Gaussian has mean of $N/2$ and standard deviation $\sqrt{N}/2$, as you can extrapolate from the case $N = 1$.) *Does it approach one of the eigenfunctions f_3 or f_4 (eqns 12.85)?*

(i) *Why did a perturbation along $f_3(x)$ not dominate the asymptotics? What symmetry forced $\epsilon_3 = 0$? Should flips of a biased coin break this symmetry?*

Using the renormalization group to demonstrate the central limit theorem might not be the most efficient route to the theorem, but it provides quantitative insights into how and why the probability distributions approach the asymptotic Gaussian form.

(12.25) **Percolation and universality.**[62] (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

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

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.[63] 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 12.5–12.6), 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[64] $\tau = 187/91$. You can test this numerically, either with the code you developed for Exercise 2.20, or by using the software at our web site [131].

(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 $p = p_c$.*[65] *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 [131]. 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 $P(p)$.[66] 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 $1+\beta\epsilon/\nu$.[67] 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

$$P(p) \sim (p_c - p)^{\beta}. \tag{12.87}$$

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 12.87 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

[63] Hence the probability that a given node is in a cluster of size S is proportional to $Sn(S)$.
[64] A non-obvious result!
[65] Conveniently, the critical probability $p_c = \frac{1}{2}$ for both these systems, see Exercise 2.20, part(c). This enormously simplifies the scaling analysis, since we do not need to estimate p_c as well as the critical exponents.
[66] 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 $L \to \infty$.
[67] We again assure the reader that these particular combinations of Greek letters are just chosen to give the conventional names for the critical exponents.

explicitly including the length L into our analysis. Let $P(p, L)$ be the mean fraction of nodes[68] 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

$$\begin{aligned} L' &= L/(1+\epsilon), \\ (p_c - p)' &= (1+\epsilon/\nu)(p_c - p), \end{aligned} \tag{12.88}$$

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 12.88). (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 12.88 and the fact that P rescales upward by $(1 + \beta\epsilon/\nu)$ under coarse-graining, write the similarity relationship for $\bar{P}$. Following our derivation of the scaling form for the avalanche size distribution (through eqn 12.11), show that $\bar{P}(L, X) = L^{-\beta/\nu}\mathcal{P}(X)$ for some function $\mathcal{P}(X)$, and hence*

$$P(p, L) \propto L^{-\beta/\nu}\mathcal{P}((p - p_c)L^{1/\nu}). \tag{12.89}$$

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 12.89 to deduce how to rescale our data. We can find the finite-sized scaling 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.[69] 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.

[68] 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.

[69] 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.

(12.26) **Hysteresis and avalanches: scaling.** (Complexity) ③

For this exercise, either download Matt Kuntz's hysteresis simulation code from the book web site [131], or make use of the software you developed in Exercise 8.17 or 8.18.

Run the simulation in two dimensions on a 1000×1000 lattice with disorder $R = 0.9$, or a three-dimensional simulation on a 100^3 lattice at $R = 2.16$.[70] The simulation is a simplified model of magnetic hysteresis, described in [130]; see also [129]. The spins s_i begin all pointing down, and flip upward as the external field H grows from minus infinity, depending on the spins of their neighbors and a local random field h_i. The flipped spins are colored as they flip, with spins in the same *avalanche* sharing the same color. An avalanche is a collection of spins which flip together, all triggered from the same original spin. The *disorder* is the ratio R of the root-mean-square width $\sqrt{\langle h_i^2 \rangle}$ to the ferromagnetic coupling J between spins:

$$R = \sqrt{\langle h^2 \rangle}/J. \tag{12.90}$$

Examine the $M(H)$ curve for our model and the $\mathrm{d}M/\mathrm{d}H$ curve. The individual avalanches should be visible on the first graph as jumps, and on the second graph as spikes. This kind of time series (a set of spikes or pulses with a broad range of sizes) we hear as *crackling noise*. You can go to our site [72] to hear the noise resulting from our model, as well as crackling noise we have assembled from crumpling paper, from fires and Rice Krispies™, and from the Earth (earthquakes in 1995, sped up to audio frequencies).

Examine the avalanche size distribution. The (unlabeled) vertical axis on the log–log plot gives the number of avalanches $D(S, R)$; the horizontal axis gives the size S (with $S = 1$ on the left-hand side). Equivalently, $D(S, R)$ is the probability distribution that a given avalanche during the simulation will have size S. The graph is created as a histogram, and the curve changes color after the first bin with zero entries (after which the data becomes much less useful, and should be ignored).

If available, examine the spin–spin correlation function $C(x, R)$. It shows a log–log plot of the probability (vertical axis) that an avalanche initiated at a point $\mathbf{x}_0$ will extend to include a spin $\mathbf{x}_1$ a distance $x = \sqrt{(\mathbf{x}_1 - \mathbf{x}_0)^2}$ away.

Two dimensions is fun to watch, but the scaling behavior is not yet understood. In three dimensions we have good evidence for scaling and criticality at a phase transition in the dynamical evolution. There is a phase transition in the dynamics at $R_c \sim 2.16$ on the three-dimensional cubic lattice. Well below R_c one large avalanche flips most of the spins. Well above R_c all avalanches are fairly small; at very high disorder each spin flips individually. The critical disorder is the point, as $L \to \infty$, where one first finds *spanning avalanches*, which extend from one side of the simulation to the other.

Simulate a 3D system at $R = R_c = 2.16$ with $L = 100$ (one million spins, or larger, if you have a fast machine). It will be fastest if you use the *sorted list* algorithm (Exercise 8.18). The display will show an $L \times L$ cross-section of the 3D avalanches. Notice that there are many tiny avalanches, and a few large ones. Below R_c you will find one large colored region forming the background for the others; this is the spanning, or *infinite* avalanche. Look at the $M(H)$ curve (the bottom half of the hysteresis loop). It has many small vertical jumps (avalanches), and one large one (corresponding to the spanning avalanche).

(a) *What fraction of the system is flipped by the one largest avalanche, in your simulation? Compare this with the hysteresis curve at $R = 2.4 > R_c$. Does it have a similar big jump, or is it continuous?*

Below R_c we get a big jump; above R_c all avalanches are small compared to the system size. If the system size were large enough, we believe the fraction of spins flipped by the spanning avalanche at R_c would go to zero. The largest avalanche would nonetheless span the system—just like the percolation cluster at p_c spans the system but occupies zero volume in the limit of large systems.

The other avalanches form a nice power-law size distribution; let us measure it carefully. Do a

[70] If you are using the brute-force algorithm, you are likely to need to run all of the three-dimensional simu system size, perhaps 50^3. If you have a fast computer, you may wish to run at a larger size, but make sure watch.

set of 10 runs (*# Runs 10*) at $L = 100$ and $R = R_c = 2.16$.[71]

Watch the avalanches. Notice that sometimes the second-largest avalanche in the view (the largest being the 'background color') is sometimes pretty small; this is often because the cross-section we view missed it. Look at the avalanche size distribution. (You can watch it as it averages over simulations.) Print it out when the simulations finish. Notice that at R_c you find a pretty good power-law distribution (a straight line on the log–log plot). We denote this critical exponent $\bar{\tau} = \tau + \sigma\beta\delta$:

$$D(S, R_c) \sim S^{-\bar{\tau}} = S^{-(\tau+\sigma\beta\delta)}. \tag{12.91}$$

(b) *From your plot, measure this exponent combination from your simulation. It should be close to two. Is your estimate larger or smaller than two?*

This power-law distribution is to magnets what the Gutenberg–Richter law (Fig. 12.3(b)) is to earthquakes. The power law stems naturally from the self-similarity.

We want to explore how the avalanche size distribution changes as we move above R_c. We will do a series of three or four runs at different values of R, and then graph the avalanche size distributions after various transformations.

Do a run at $R = 6$ and $R = 4$ with $L = 100$, and make sure your data files are properly output. Do runs at $R = 3$, $R = 2.5$, and $R = 2.16$ at $L = 200$.

(c) *Copy and edit your avalanche size distribution files, removing the data after the first bin with zero avalanches in it. Start up a graphics program, and plot the curves on a log–log plot; they should look like power laws for small S, and cut off exponentially at larger S. Enclose a copy of your plot.*

We expect the avalanche size distribution to have the scaling form

$$D(S, R) = S^{-(\tau+\sigma\beta\delta)}\mathcal{D}(S(R-R_c)^{1/\sigma}) \tag{12.92}$$

sufficiently close to R_c. This reflects the similarity of the system to itself at a different set of parameters; a system at $2(R - R_c)$ has the same distribution as a system at $R - R_c$ except for an overall change A in probability and B in the size scale of the avalanches, so $D(S, R - R_c) \approx AD(BS, 2(R - R_c))$.

(d) *What are A and B in this equation for the scaling form given by eqn 12.92?*

At $R = 4$ and 6 we should expect substantial corrections! Let us see how well the collapse works anyhow.

(e) *Multiply the vertical axis of each curve by $S^{\tau+\sigma\beta\delta}$. This then should give four curves $\mathcal{D}(S(R - R_c)^{1/\sigma})$ which are (on a log–log plot) roughly the same shape, just shifted sideways horizontally (rescaled in S by the typical largest avalanche size, proportional to $1/(R - R_c)^{1/\sigma}$). Measure the peak of each curve. Make a table with columns R, S_{peak}, and $R - R_c$ (with $R_c \sim 2.16$). Do a log–log plot of $R - R_c$ versus S_{peak}, and estimate σ in the expected power law $S_{\text{peak}} \sim (R - R_c)^{-1/\sigma}$.*

(f) *Do a scaling collapse: plot $S^{\tau+\sigma\beta\delta}D(S, R)$ versus $(R - R_c)^{1/\sigma}\, S$ for the avalanche size distributions with $R > R_c$. How well do they collapse onto a single curve?*

The collapses become compelling only near R_c, where you need very large systems to get good curves.

[71] If your machine is slow, do fewer. If your machine is fast, use a larger system. Make sure you do not run out of RAM, though (lots of noise from your hard disk swapping); if you do, shift to the *bits* algorithm if its available. *Bits* will use much less memory for large simulations, and will start up faster than *sorted list*, but it will take a long time searching for the last few spins. Both are much faster than the brute-force method.