1 Universal Exponents and Scaling Functions from the Renormalization Group

A renormalization group transformation involves a *coarse graining* operation, followed by a *rescaling* of various amplitudes. It yields quantitative predictions when we assume that repeated transformations lead to a fixed point (which is then self-similar), or when different starting points can be shown to flow together under coarse-graining (thus sharing common behavior on long scales). In these notes, we will argue that, for ordinary renormalization-group fixed points (i.e., allowing continuous rescaling and without marginal eigenvalues) a long-wavelength function A depending on N variables $\theta_1, \ldots \theta_N$ can be written as a power law times a function of $N-1$ new variables[1]

$$A(\theta_1, \ldots, \theta_N) = \theta_1^{X_1} \mathcal{A}(\theta_2/\theta_1^{x_2}, \ldots, \theta_N/\theta_1^{x_N}). \tag{1}$$

Here X_1 is a universal critical exponent, as are the exponents $x_2, \ldots x_N$ used to create the new combined variables, and $\mathcal{A}$ is a universal scaling function.[2]

We can make the power law outside that of any variable we wish by permutation. The combinations of variables used as arguments inside the scaling function, we will see, are invariant under the renormalization group; we may use any independent $N-1$ invariant combinations of variables to form our scaling function. So, for example, the spin-spin correlation function $C(r,t|H,T) = \langle s(x,t_0)s(x+r,t_0+t)\rangle$ of a magnet in a field H and temperature T can be written in the scaling forms

$$\begin{aligned} C(r,t|H,T) &= r^{-(2-d+\eta)}\tilde{\mathcal{C}}(t/r^{z/\nu}, |T-T_c|/r^{-1/\nu}, H/r^{-\beta\delta/\nu}) \\ &= r^{-(2-d+\eta)}\mathcal{C}(r/|T-T_c|^{-\nu}, t/|T-T_c|^{-z}, H/|T-T_c|^{\beta\delta}) \end{aligned} \tag{2}$$

where the second form $\mathcal{C}(a,b,c) = \tilde{\mathcal{C}}(b/a^{z/\nu}, a^{1/\nu}, c\,a^{\beta\delta/\nu})$ is more useful.

Rather than derive this formula in the abstract, we'll work with three examples: the avalanche size distribution $D(S) \sim S^{-\tau}$ at a critical point, the magnetization and correlation lengths $M(T) \sim (T_c - T)^\beta$ as we approach the Ising model critical point at $t = T - T_c = 0$, and the cluster size distribution $D(S|\,p)$ near the percolation critical point at $p = P - P_c = 0$.

[1] There is a nice analogy between scale invariance and translation invariance. In a translation invariant mechanical system of N particles (without friction) we can move to a center-of-mass frame, solving an N-dimensional problem in terms of functions of $N-1$ relative coordinates.

[2] We're oversimplifying a bit here. There should be arbitrary constants rescaling the various factors, so $A(\theta_1, \ldots, \theta_N) = A_1\theta_1^{X_1}\mathcal{A}(a_2\theta_2/\theta_1^{x_2}, \ldots, a_N\theta_N/\theta_1^{x_N})$, with the factors A_1, a_2, ... corresponding to the rescalings A and B of density and temperature in the liquid-gas collapse of Fig. 12.6. Also, there can be analytic corrections to scaling (the tilting of the singular diameter in the same figure), singular corrections to scaling, etc.

1.1 Avalanche Size Distribution: Power laws at Critical Points

We have seen that both earthquakes and crackling noise in magnets share a power law distribution of event sizes $D(S) \sim S^{-\tau}$. Imagine coarse-graining a system of avalanches, measuring the avalanches with a new ruler with units a factor b larger than the previous, and ignoring avalanches smaller than length b. (For example, $b = 3$ in the Ising majority-rule coarse-graining shown in figure 12.9.) One can see that the avalanche sizes (i.e., volumes) will, in the new units, be a factor of b^{d_f} smaller, $S \to S/b^{d_f}$. Here one might imagine, for three-dimensional systems, that d_f would be equal to three. But for many avalanche models, the avalanches are *fractal* – with holes filling almost the entire volume. (The percolation clusters in figures 12.2 and 12.7 are fractal, with net density zero in the limit of infinite system size.) If d_f equals the system dimension, the avalanches are said to be *compact*; otherwise d_f is their fractal dimension.

Next we need to rescale the avalanche size distribution. The new probability density evaluated at S/b^{d_f} will be the old one evaluated at S, but multiplied by a factor b^a for some constant a: $D' = b^a D$. This rescaling is necessary to keep the probability density normalized, and has two different contributions. First, the avalanches in the range $[S, S + \mathrm{d}S]$ get coarse-grained into new range that's smaller by b^{d_f}, so the density must go up by that factor. Second, some of the smaller avalanches are now below the new cutoff, and so don't contribute to the new normalization. In the end, $D'(S/b^{d_f}) = b^a D(S)$, and our renormalization-group transformation[3] is

$$T[D](S) = b^a D(S\, b^{d_f}). \tag{3}$$

[3]Our notation here is sloppy. The renormalization group really acts on the system, not on the size distribution D. If the rules for our system are specified by $\mathcal{H}$, then eqn 3 should really describe the avalanche size distribution of the renormalized system $D[T[\mathcal{H}]](S) = b^a D[\mathcal{H}](S\, b^{d_f})$.

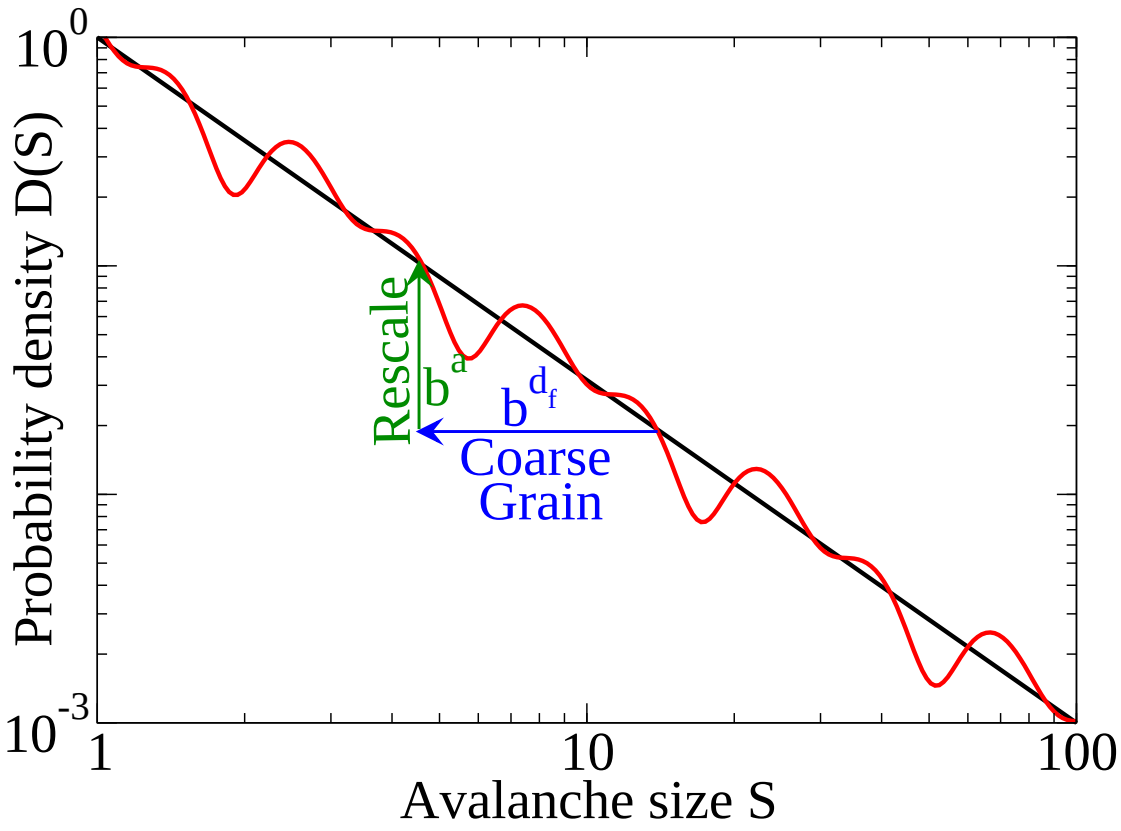

Fig. 1 Power laws are the only scale-invariant function, if the scale invariance is continuous. Discrete scale invariance can lead to non-power-law behavior.

To extract predictions, we must posit that our physical system converges under our renormalization group to a fixed point D^*, so $T[D^*] = D^*$ and $D^*(S) = b^a D^*(S\,b^{d_f})$. This tells us the probability density $D^*(S)$ at size S in terms of the probability density for a larger avalanche of size $S\,b^{d_f}$ (manifesting the scale invariance). Suppose, for each size S, we repeat the transformation n times, until we reach some large size $S\,b^{nd_f} = S_0$ (say a billion). Now we know that $b^{-d_f} = S/S_0$, so $b^a = (S/S_0)^{-a/d_f}$, and hence

$$D(S) = (S/S_0)^{-a/d_f} D(S_0) = S^{-\tau}[S_0^{\tau} D(S_0)] \sim S^{-\tau} \tag{4}$$

where we define $\tau = -a/d_f$. This general procedure leads to power laws for many functions of one variable at critical points. Because many systems flow to the same fixed point, at large sizes their avalanches will all share the same value of τ – the critical exponent τ is universal.

Our derivation in the last paragraph is somewhat more complicated than it really had to be. Basically, what we are arguing is that power laws are the only scale invariant function. Think of plotting $D^*(S)$ on a log-log plot, as in Fig. 1. By being a renormalization group fixed point, the curve must land onto itself when shifted to the left by $\log(b^{d_f})$ and upward by $\log b^a$, suggesting a straight line (power law) with the form $D(S) \sim S^{-a/d_f}$. In this

construction, we see that we can also find renormalization-group fixed points that are power laws times periodic functions in $\log S$ with period $\log b^{d_f}$ (curvy line in Fig. 1). In systems like Feigenbaum period doubling, where the coarse-graining factor b is not arbitrary (you need to coarse grain by powers of two to get similar behavior) one does indeed find scaling functions that show these features (see Fig. 12.16). But for many systems, the factor b can be chosen as infinitesimal, and power laws are then implied by scale invariance.

1.2 Magnetization and correlation length: Power Laws near Criticality

Power laws also arise for scalar quantities *near* critical points. Figure 2 shows the magnetization M and the correlation length ξ of a three-dimensional Ising model near T_c. How do we explain these power laws using the renormalization group? Systems with $T \neq T_c$ don't flow to the fixed point! ξ and M measure *violations* of self-similarity.

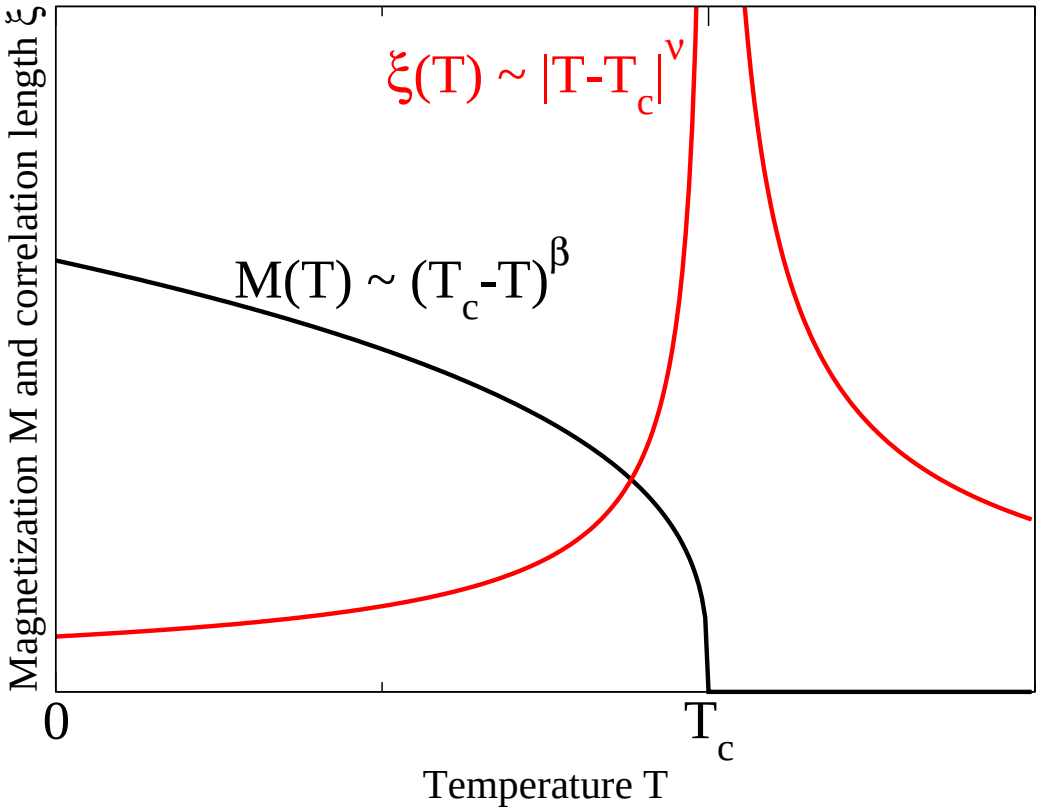

Fig. 2 Power laws near criticality, magnetization and correlation length near T_c for the 3D Ising model.

Figure 3 shows how systems starting near a critical point flow under the renormalization

group. The key is that the systems close to T_c first flow very near to the fixed point, and then all depart from the critical region along the same line.

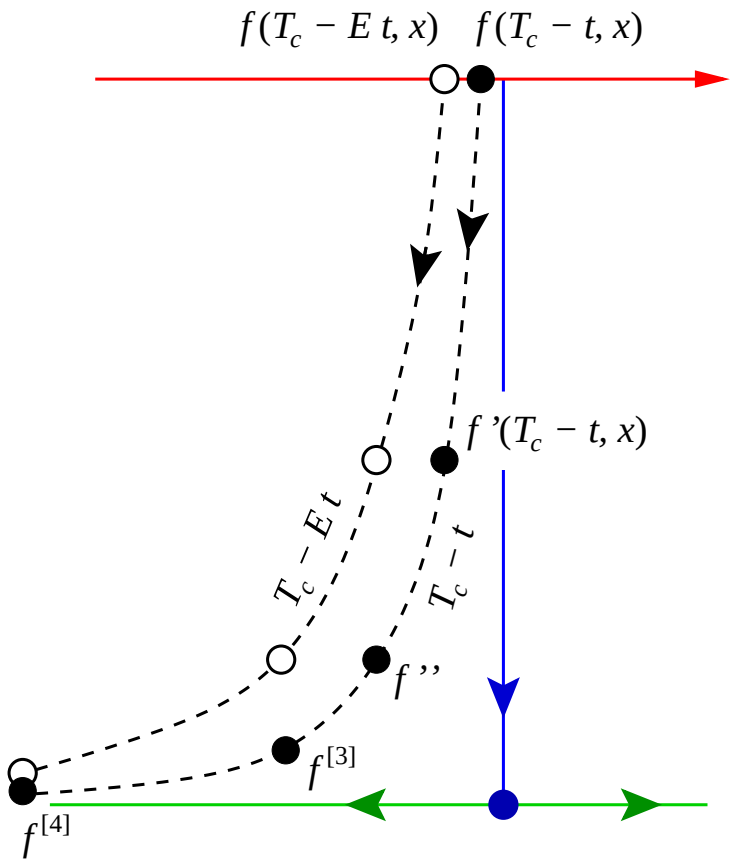

Fig. 3 Scaling near criticality. If two points in system space flow towards one another under coarse-graining, their behavior must be similar on long length scales. Here we measure a function $f(x)$ for our system (top line) at two different temperatures, $T_c - t$ and $T_c - Et$. The dots represent successive coarse-grainings by a factor B; under this renormalization group $f \to f' \to f'' \to f^{[3]} \ldots$. Here $f(T_c - t, x)$ after four coarse-grainings maps to nearly the same system as $f(T_c - Et, x)$ after three coarse-grainings. We thus know, on long length scales, that $f'(T_c - t, x)$ must agree with $f(T_c - Et, x)$; the system is similar to itself *at a different set of external parameters*. In particular, each coarse-graining step changes x by a factor B and f by some factor A, so $f'(T_c - t, y) = Af(T_c - t, By) = f(T_c - Et, y)$ for large distances y.

Note that the (two-dimensional) figure shows a stable line of points flowing downward into the fixed point, and an unstable line horizontally flowing away. Remember that this is a cartoon of a flow in an infinite dimensional space. If there is only one relevant variable (like temperature here), we understood that the systems which flow into the fixed point form a surface (the *stable manifold*) with one fewer dimensions than the space (co-dimension one).

In contrast, the *unstable manifold* shown here as the horizontal line is really just a one-dimensional curve. Hence if a system at two different original temperatures both flow out along the same curve, then they look the same as each other at different amounts of coarse-graining. Thus, in the figure, the quantity f at $T_c - t$ after four coarse-grainings agrees with f at $T_c - Et$ after three coarse-grainings: the coarsened system is similar to itself *at a different temperature*.

To what new reduced temperature Et should our coarse-grained system be similar to? The new temperature must have a shorter correlation length, so it will be farther from the critical temperature (so $E > 1$) and farther from the critical surface (by the same factor[4] E). As the two systems are coarse-grained, they will first flow into the proximity of the fixed point, with some smooth distortions that will typically stretch their distances to the critical surface. In the limit $t \to 0$, the two points will end up after several coarse-grainings (say m) lined up along the unstable manifold (the green line), with $T^m[f](Et)$ a factor of E farther out than $T^m[f](t)$.

Now, locally near the fixed point the renormalization-group transformation can be linearized ($\nabla T[f] \cdot \delta \sim T[f+\delta] - T[f]$, a kind of functional Jacobian describing the local flows around the fixed point). Our assumption that there is one unstable direction is the same as the assumption that our Jacobian has one eigenvalue λ_0 greater than one. The unstable manifold departs from the fixed point along the corresponding eigenvector.[5] If we choose $E = b^{1/\nu} = \lambda_0$, then after one extra coarse-graining the system nearer to the critical point will land atop the farther one:

$$T^{m+1}[f](t) = T^m[f](b^{1/\nu}t). \tag{5}$$

Let's start with the correlation length ξ of our system. This is often described by the decay of the correlation function at long distances, but you can view it as the typical size of the largest clusters of up spins. ξ is known to diverge at T_c like $|t|^\nu$, with[6] $t = (T - T_c)$. Since ξ is a length, when we coarse graining to change our unit of length by a factor of b, it changes simply as $\xi \to \xi/b$. Using eq 5, we find $\xi(t)/b^{m+1} = \xi(b^{1/\nu})/b^m$, so

$$\xi(t) = b\xi(b^{1/\nu}t). \tag{6}$$

Renormalizing n times until $b^{n/\nu}t = 1$, we find $b^n = t^{-\nu}$, and

$$\xi(t) = t^{-\nu}\xi(1) \sim t^{-\nu} \tag{7}$$

justifying after the fact our choice of $b^{1/\nu}$ as the right way to write the unstable eigenvector.

[4]This is true so long as the physical system (top line in Fig. 3) crosses the critical surface at a non-zero angle (i.e., *transversally*).

[5]We will see later that J is not symmetric, so we will need to be careful about whether this is a left or right eigenvector.

[6]The *reduced temperature* t is traditionally defined to be $(T - T_c)/T_c$. The normalization of T_c in the denominator is simply a convenient way to nondimensionalize it. We ignore it here, for reasons similar to our ignoring A_1 and $a_2, a_3, \ldots$ in footnote 2.

Now let's consider the scaling behavior of the magnetization shown in Fig. 2. When we coarse grain by a length factor b, M will need to be rescaled by some factor which we may write as $b^{\beta/\nu}$, so $T[M](t) = b^{\beta/\nu} M(b^{1/\nu} t)$. Iterating $m+1$ times at t and m times at $Et = b^{1/\nu} t$ as in equation 5, we have $T^{m+1}[M](t) = M(t)\, b^{(m+1)\beta/\nu} = T^m[M](Et) = M(\beta^{-1/\nu} t)\, b^{m\beta/\nu}$ (that is, the two systems coarse-grain to the same trajectory leaving the fixed point), so we get a scaling relation $M(t) = b^{-\beta/\nu} M(b^{-1/\nu} t)$. Iterating this n times until $b^{-n/\nu} t = -t_0$ (remembering that $T < T_c$ so $t = T - T_c < 0$), we find $b^{-n\beta/\nu} = (-t/t_0)^\beta = |t/t_0|^\beta$ and hence

$$M(t) = |t|^\beta \left(M(-t_0)/t_0^\beta \right) \propto |t|^\beta = (T_c - T)^\beta, \tag{8}$$

giving us the power-law dependence of the magnetization on the distance to the critical temperature.

1.3 Cluster size distribution, and universal scaling functions

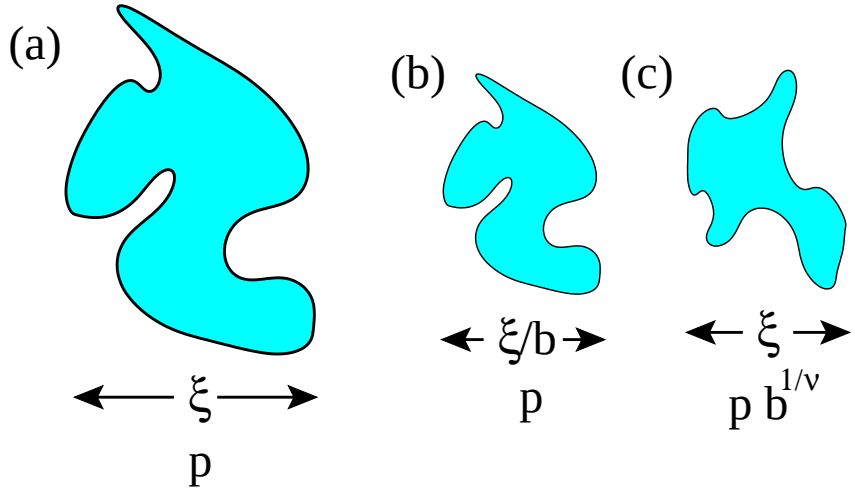

Fig. 4 Coarse-graining and rescaling of a percolation cluster.

To illustrate the derivation of scaling laws for functions of more than one variable, let us consider the distribution $D(S, p)$ of percolation cluster sizes (Fig. 5) as a function of distance $p = P - P_c$ to the critical point.[7] Using the scaling of D and S for the avalanche size distribution (eqn 3), we get the renormalization-group transformation

$$T[D](S|\, p) = b^a D(S\, b^{d_f} |\, p). \tag{9}$$

That is, under coarse-graining S shifts to S/b^{d_f}, and then to keep the probability distribution normalized we must then multiply by a factor b^a (see Fig. 5).

[7] Remember that P is the probability that a site or a bond exists (has not been punched out), see Fig. 12.2.

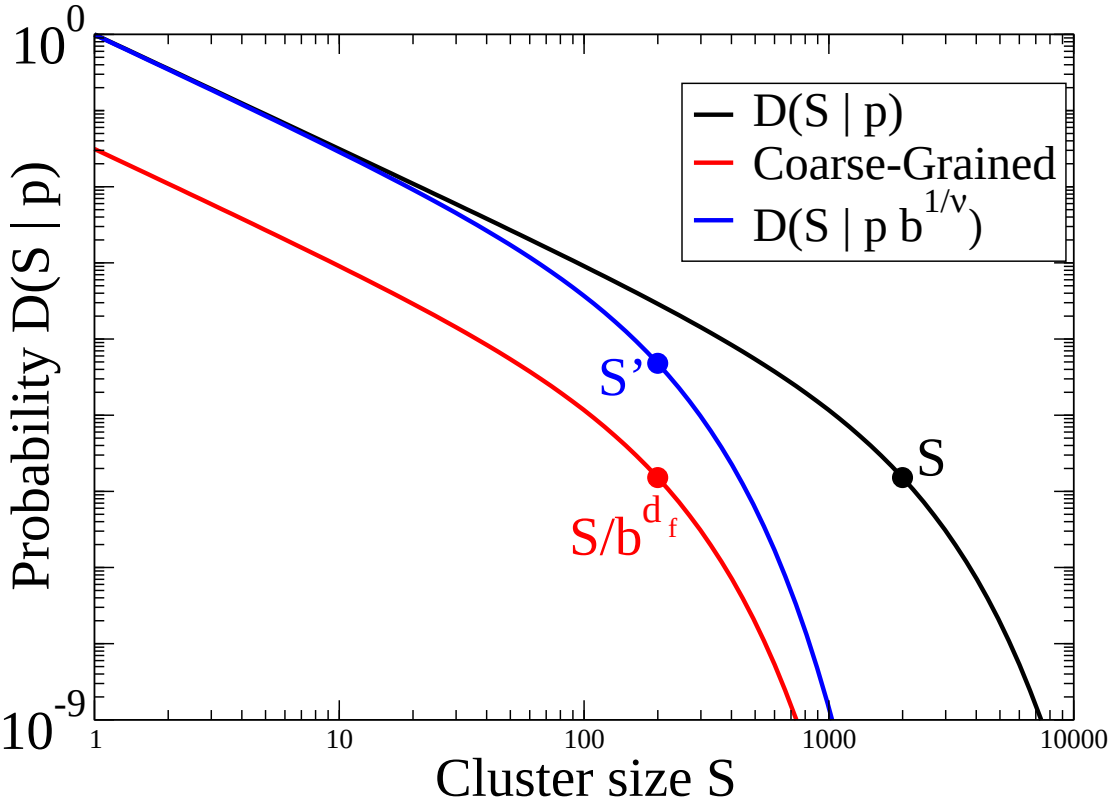

Fig. 5 Coarse-graining and rescaling of a cluster size distribution. Notice that the three curves are the same shape, just shifted sideways and/or upward, despite appearances to the contrary.

Since we are close to the critical point, we again may assume that after m rescalings to approach the critical point (as in eqn 5) $T^{m+1}[D](p) = T^m[D](b^{1/\nu}t)$[8] and hence

$$b^{(m+1)a} D(S_0 b^{(m+1)d_f} | \, p) = b^{ma} D(S_0 b^{md_f} | \, p \, b^{1/\nu}, \tag{10}$$

implying for large clusters of size $S = S_0 b^{(m+1)d_f}$ that

$$D(S| \, p) = b^a D(S/b^{d_f} | \, p \, b^{1/\nu}). \tag{11}$$

Iterating n times until $|p| \, b^{n/\nu} = 1$, we have $b^{na} = |p|^{-\nu a}$, $b^{n \, d_f} = |p|^{-\nu d_f}$, and

$$\begin{aligned} D(S| \, p) &= b^{na} D(S/b^{n \, d_f} | \pm 1) = |p|^{-a\nu} D(S/|p|^{-\nu d_f} | \pm 1) \\ &= |p|^{-a\nu} \tilde{\mathcal{D}}_{\pm}(S/|p|^{-\nu d_f}). = |p|^{-a\nu} \tilde{\mathcal{D}}_{\pm}(S|p|^{\sigma}). \end{aligned} \tag{12}$$

where we define the maximum cluster size exponent $\sigma = \nu d_f$, so $S \sim p^{-\sigma}$ gives the 'knee' in the avalanche size distribution in Fig. 5. Here $\tilde{\mathcal{D}}$ is a *universal scaling function*, with $\tilde{\mathcal{D}}_+$

[8]Here the correlation length exponent ν is different from that for the Ising model, of course, since the unstable eigenvalue $b^{1/\nu}$ is now that for the percolation fixed point, not the Ising fixed point.

giving the shape of the cluster size distributions above P_c (the curves in Fig. 5) and $\tilde{\mathcal{D}}_-$ giving the shape below P_c.

This is not the traditional form in which to express the scaling function; the power law $S^{-\tau}$ is hidden as the dependence of $\tilde{\mathcal{D}}$ at small arguments. We change from a power of p outside to a power of S by pulling out the appropriate power of the argument:

$$\begin{aligned} D(S|\,p) &= |p|^{a\nu}(S|p|^\sigma)^{-a\nu/\sigma}\left[(S|p|^\sigma)^{-a\nu/\sigma}\tilde{\mathcal{D}}_\pm(S|p|^\sigma)\right] \\ &= S^{-a\nu/\sigma}\mathcal{D}_\pm(S|p|^\sigma) \\ &= S^{-\tau}\mathcal{D}_\pm(S|p|^\sigma) \end{aligned} \tag{13}$$

where we recognize $\tau = a\nu/\sigma = a\nu/(\nu d_f) = a/d_f$ from eqn 4. This form of the scaling function will usually go to a constant at small values of its argument,[9] hence giving us $D(S|p=0) = S^{-\tau}\mathcal{D}(0) \sim S^{-\tau}$ at the critical point $P = P_c$, as we saw before.

[9] If not, it's said to be a *singular scaling function*.