Wednesday

Read: Chapter 12, Sec. 12.1 (Universality)

1. Renormalization-group trajectories. ①

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 $t' = (1 + ae) t$ and a shifted external field $h' = (1 + be) h$, with $\epsilon$ 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 supposedly 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?)

\[ \text{(A)} \]

\[ \text{(B)} \]
The solid dots are at temperature $t_0$; the open circles are at temperature $t = 2t_0$.

(b) In terms of $\epsilon$ 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 $(1 - \epsilon)^n \approx \exp(-ne)$ 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 $\hat{h}$ for the corresponding open circle, in terms of $a$, $b$, $\epsilon$, and the original coordinates? (You may use the relation between $\hat{h}$ 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 $M'(1 + ce)M$, so $M((1 + ae)t, (1 + be)h) = (1 + ce)M(t, h)$.

(c) Suppose $M(t, h)$ is known at $(t_0, h_0)$, one of the solid dots. Give a formula for $M(2t_0, \hat{h})$ at the corresponding open circle, in terms of $M(t_0, h_0)$, the original coordinates, $a$, $b$, $c$, and $\epsilon$. (Hint: Again, rescale $n$ times.) Substitute your formula for $\hat{h}$ 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 $2 \rightarrow 1/t_0$ in your formula from part (c). Show that $M(t, h) = t^\beta M(h/t^{\delta})$ (the standard scaling form for the magnetization in the Ising model). What are $\beta$ and $\delta$ in terms of $a$, $b$, and $c$? How is $M$ related to $M(t, h)$ where the curve crosses $t = 1$?

Friday
Read: Chapter 12, Sec. 12.2 (Scale Invariance)

2. Scaling and coarsening. (Condensed matter) ③

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 $C_{\text{coar}}^{t_0}(r, T) = c(r)$.

Assume that the correlation function at short distances $C_{\text{coar}}^t(0, T, t)$ 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 $t_0$, $C_{\text{coar}}^{2t_0}(r, T)$, in terms of $c(r)$. Write a scaling form

$$C_{\text{coar}}^t(r, T) = t^{-\omega} C(r/t^\rho, T). \quad (1)$$

Use the time independence of $C_{\text{coar}}^t(0, T)$ and the fact that the order parameter is not conserved (Section 11.4.1) to predict the numerical values of the exponents $\omega$ and $\rho$.

It was only recently made clear that the scaling function $C$ 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.

Monday
Read: Chapter 12, Sec. 12.3 (Examples of critical points)
(Pre-class question longer than usual!)
3. **Scaling and corrections to scaling.** (Condensed matter) ②

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

$$\chi(T) = A(T - T_c)^{-1.25} + B + C(T - T_c) + D(T - T_c)^{1.77}. \tag{2}$$

(a) *Assuming this is the correct dependence near $T_c$, what is the critical exponent $\gamma$?*

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

$$C(r, T) = r^{-0.026}C(r(T - T_c)^{0.65}), \tag{3}$$

where the function $C(x)$ is found to be roughly $\exp(-x)$.

(b) *What is the critical exponent $\nu$? The exponent $\eta$?*

4. **Period Doubling.** (Scaling) ③

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).

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

(a) From the diagram shown, roughly estimate the values of the Feigenbaum numbers $\delta$ (governing the rescaling of $\mu - \mu_\infty$) and $\alpha$ (governing 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 $\zeta$ and $\nu$ 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 $\zeta$ in terms of $\alpha$ and $\delta$. If $\xi$ 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 $\nu$ in terms of $\alpha$ and $\delta$? (Hint: be sure to check the signs.)

**Exercises**

12.10 or 12.11: The renormalization group and the central limit theorem (short, or long version).

1. **Diffusion equation and universal scaling functions**

   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 $\rho(x, t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left(-\frac{(x - vt)^2}{4Dt}\right)$. 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 $t' = t/2, x' = x/\sqrt{2}, \rho' = \sqrt{2}\rho$. These three together constitute our renormalization group operation.

   (a) Write an expression for $\rho'(x', t')$ in terms of $D, v, x', t'$ (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, $X = x/\sqrt{t}$ and $V = \sqrt{t}v$. We can then write

   $$\rho(x, t) = t^{-\alpha}P(X, V, D), \quad (4)$$

   ![This exercise was developed in collaboration with Archishman Raju.](https://example.com/1)

   ![Because $\rho$ is a density, we need to rescale $\rho' dx' = \rho dx$.](https://example.com/2)
a power law times a universal scaling function of invariant combination of variables.

(b) Show that $X$ and $V$ are invariant under our renormalization group operation. What is $\alpha$? Write an expression for $\mathcal{P}$, in terms of $X$, $V$, 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 $\alpha$ and the invariant scaling combinations $X$, $V$, and $D$ are determined just by the coarsening and rescaling 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 (4) is expected.)

2. Biggest of bunch: Gumbel. (Mathematics, Statistics, Engineering)

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

$$F_1(H) = P(H' < H) = \int_{H'}^{H} \rho_1(H') \, dH'.$$

(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. 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

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

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

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

for large $H$. (Why is the probability $F_N(H)$ small when $H$ is not large, at large $N$?) What are $\mu$ and $\beta$ for this case?
The constants $\beta$ and $\mu$ 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

$$
F(x) = \exp(-\exp(-x)) \\
\rho(x) = \partial F/\partial x = \exp(-(x + \exp(-x))).
$$

(7)

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\[^{[3],\text{section 8.3}}\]. In particular, suppose

$$
F_1(H) \approx 1 - A \exp(-BH^\delta)
$$

(8)
as $H \to \infty$ for some positive constants $A$, $B$, and $\delta$. 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 5).

(c) Show that the extreme value statistics $F_N(H)$ for this distribution is of the Gumbel form (eqn 6) 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\[^{[3]}\]. (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 $F_1(H) = (1/2)(1 + \text{erf}(H/\sqrt{2}))$ 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 8 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 \[^{[3]}\].

(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 $\rho_N(H) = dF_N/dH$ 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 A.6). Distributions with power-law tails (like the distributions

\[^{3}\text{The Gumbel distribution can also describe extreme values for a bounded distribution, if the probability density at the boundary goes to zero faster than a power law.}\]
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 have extreme value distributions that are described by Weibull statistics (see Exercise 3).

3. **First to fail: Weibull** (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 2 and A.6), 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 \(S(t) = \int_t^\infty \rho(t') dt'\) are often given by the Weibull distribution

\[
S(t) = e^{-(t/\alpha)^\gamma},
\]

\[
\rho(t) = \frac{dS}{dt} = \frac{\gamma}{\alpha} \left( \frac{t}{\alpha} \right)^{\gamma-1} e^{-(t/\alpha)^\gamma}.
\] ⑨

Let us begin by assuming that the processors have a constant rate \(\Gamma\) 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 \(\epsilon\), show that the the probability \(S_N(t)\) at time \(t\) that all \(N\) processors are still running is of the Weibull form (eqn 9). What are \(\alpha\) and \(\gamma\)?

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

\[
\rho_1(t) \sim Bt^k
\]

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 10) is of the Weibull form for large \(N\), and give \(\alpha\) and \(\gamma\) as a function of \(B\) and \(k\).

The parameter \(\alpha\) in the Weibull distribution just sets the scale or units for the variable \(t\); only the exponent \(\gamma\) really changes the shape of the distribution. Thus the form of

\[4\] More specifically, bounded distributions that have power-law asymptotics have Weibull statistics; see note 3 and Exercise 3 part (d).

\[5\] Developed with the assistance of Paul (Wash) Wawrzynek.
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 \textit{universality} 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.

\textit{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 \( \sigma \) the material can support before it will snap. Fracture in brittle materials often occurs due to pre-existing microcracks, typically on the surface of the material. Suppose we have an isolated 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 given by

\[ \sigma_c(L) \approx K_c/\sqrt{\pi L}. \]  

(11)

Here \( K_c \) is the \textit{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 \( \sigma \) at their tips into a square-root singularity; longer cracks have more stress to concentrate, leading to eqn 11.)

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

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

\footnote{The Weibull distribution forms a one-parameter family of universality classes; see chapter 12.}

\footnote{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 \textit{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.}

\footnote{The interactions between microcracks are often not small, and are a popular research topic.}

\footnote{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...}
not yield a Weibull distribution, since the probability of a small critical stress does not vanish as a power law $B\sigma^k$ (eqn 10).

(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 $\sigma$). (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, 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 $\eta$?

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.