Pre-class Preparation

Wednesday
Read: Chapter 3, Sec. 3.3 (Temperature) and 3.4 (Pressure & Chemical Potential; 3.4.1 is optional), Sec. 3.5 (Entropy & fussy stuff).

1. **Entropy maximization and temperature.**

   Explain in words why, for two weakly coupled systems, that eqn 3.23
   \[
   \rho(E_1) = \Omega_1(E_1)\Omega_2(E_2)/\Omega(E) \tag{1}
   \]
   should be intuitively clear for an equilibrium system. Using \( S = k_B \log(\Omega) \), show in one step that maximizing the probability of \( E_1 \) maximizes the sum of the entropies \( S(E_1) + S(E_2) \). Show in one more step that \( 1/T = \partial S/\partial E \) is equal for the two systems when the probability is maximized.

2. **What is the chemical potential?**

   What the heck is a chemical potential? Explain.

Friday
Read: Chapter 4, Sec. 4.1 (Liouville’s theorem) Sec. 4.2 (Ergodicity)

3. **Undistinguished particles.**

   Why should we divide the phase-space volume by \( N! \), when we do not keep track of the differences between the \( N \) particles? Look up ‘entropy of mixing’. In Fig. (5.4), can you see how to extract work from the mixing using methods that cannot tell white from black particles? If we had a door in the partition wall that let through only white particles, what would happen? Could we then extract work from the system?
Monday

Read: Chapter 4, Sec. 4.2 (Ergodicity)

4. **Perverse initial conditions.**

If we start a gas of classical spherical particles in a square box all in a vertical line, and all moving exactly in the vertical direction, they will bounce back and forth vertically forever.

*Does that mean a gas of particles in a square box is not ergodic? Why or why not?*

**Exercises**

7.3: *Phase space units and the zero of entropy* (if we don’t finish it in class).

3.5: *Hard sphere gas* (if we don’t finish it in class).

A.19: 2D turbulence and Jupiter’s great red spot. (hints available)

5. **2D turbulence and Jupiter’s Great Red Spot** ([Astrophysics])

Fully-developed turbulence is one of the outstanding challenges in science. The agitated motion of water spurting through a fire hose has a complex, fluctuating pattern of swirls and eddies spanning many length scales. Many have studied the velocity-velocity correlation functions in turbulence, and there are close analogies to the scaling behavior seen at continuous phase transitions. Nobody yet, though, has found the links to the renormalization group.

Turbulence in two-dimensional fluids is much better understood, with the latest fashion using conformal field theories (see A. M. Polyakov, *Nucl. Phys. B* **396** (1993) and much further work), but the original insight was a point vortex model introduced by Onsager and others that we will simulate here.

The model describes the motion of point *vortices*, describing local rotational eddies in the fluid. It provides an analogy to Liouville’s theorem, an example of a system with an infinite number of conserved quantities, and a system which can have a negative temperature. And it provides a plausible explanation of Jupiter’s Giant Red Spot, and cyclones, hurricanes, and typhoons on Earth. We draw on many sources, but our inspiration came from the work of Jonathan Miller, Peter B. Weichman, and M. C. Cross, (“Statistical mechanics, Euler’s equation, and Jupiter’s Red Spot”, *Phys. Rev. A* **45**, 2328 (1992)).

---

1 This exercise was developed in collaboration with Jaron Kent-Dobias.

2 Cyclones, hurricanes, and typhoons are different names used for the same weather patterns.
Fig. 1 Jupiter’s Red Spot. Jupiter’s atmosphere is stormy, with winds that can exceed 400 miles per hour in a complex, turbulent pattern. Jupiter’s ‘red spot’ is a giant storm 3 1/2 times the size of the earth, that has lasted for at least 186 years. (Image from Voyager I, NASA’s Goddard Space Flight Center, courtesy of NASA/JPL, [http://www.nasa.gov/content/jupiters-great-red-spot-viewed-by-voyager-i](http://www.nasa.gov/content/jupiters-great-red-spot-viewed-by-voyager-i).)

The flow of most liquids is almost incompressible; if \( u(x) \) is the velocity field, then to an excellent approximation \( \nabla \cdot u = 0 \). Helmholtz’s theorem tells us that a smooth field \( u \) is determined by its divergence and curl, if it dies away fast enough at infinity and the region has no holes.\(^3\) So knowing how the vorticity \( \omega = \nabla \times u \) evolves is enough to determine the evolution of the fluid. In two dimensions, the vorticity is a scalar \( \omega(x) = \partial u_y/\partial x - \partial u_x/\partial y \).

(a) Show that

\[
 u_0(x) = \frac{\Gamma}{2\pi} \frac{\hat{\theta}}{r}
\]

in two dimensions has both zero curl and zero divergence, except at zero. (This is the velocity field around one of Onsager’s vortices.) Use Stokes’ theorem to show that the curl of this velocity field is \( \omega(x) = \Gamma \delta(x) = \Gamma \delta(x) \delta(y) \). Argue using this solution that the velocity field in an annulus is not uniquely defined by its divergence and curl alone (as described by footnote\(^3\)). (Hint: Consider the annulus bounded by two concentric circles surrounding our vortex.)

We provide a simulation of the dynamics of interacting vortices. Download the hints file.

\(^3\) More generally, if the vorticity is defined in a non-simply connected region \( S \) with \( n \) holes, and if \( u \) is parallel to the boundaries, then knowing \( \omega(r) \) and the circulations \( \int_{\partial S_n} u \cdot d\ell \) around the edges \( \partial S_n \) of the holes specifies the velocity \( u \) inside.
(b) Run the simulation with one vortex with $\Gamma_1 = 1$ and a ‘tracer vortex’ with $\Gamma = 0$. Does the tracer vortex rotate around the test vortex?

At low velocities, high viscosity, and small distances, fluids behave smoothly as they move; turbulence happens at high Reynolds numbers, where velocities are big, distances small, and viscosities are low. If we reduce the viscosity to zero, the kinetic energy of the fluid is conserved.

We can write $u$ as a nonlocal function of $\omega$

$$u_x(r) = -\frac{1}{2\pi} \int dr' \frac{y - y'}{(r - r')^2}\omega(r') \quad (3)$$

$$u_y(r) = \frac{1}{2\pi} \int dr' \frac{x - x'}{(r - r')^2}\omega(r') \quad (4)$$

(the Biot-Savart law).

(c) Show that the Biot-Savart law agrees with your answer from part (a) for the case of a $\delta$-function vorticity.

We can write the kinetic energy $H$ in terms of the vorticity, which turns out to be a non-local convolution

$$H = \int \frac{1}{2}u(r)^2dr$$

$$= -\frac{1}{4\pi} \int dr \int dr'\omega(r)\omega(r') \log |r-r'|$$

where we set the density of the fluid equal to one and ignore an overall constant.\(^4\) We can also write the equations of motion for the vorticity $\omega$ in terms of itself and the velocity $u$

$$\frac{\partial \omega}{\partial t} = -u \cdot \nabla \omega, \quad (6)$$

To set up our simulation, we need to find a numerical representation for the vorticity field. Onsager suggested a discretization into vortices (part (a)), $\omega(r) = \sum_{i=1}^{N} \Gamma_i \delta(r - r_i)$. For point vortices, the energy becomes

$$\mathcal{H} = -\frac{1}{2\pi} \sum_i \sum_{j \neq i} \Gamma_i \Gamma_j \log \left( (r_i - r_j)^2 \right). \quad (7)$$

Note that there is no ‘kinetic energy’ for the vortices. The energy in the velocity field is purely kinetic energy; the ‘potential’ energy of the vortices is the kinetic energy of the fluid.

\(^4\)Miller, Weichman, and Cross tell us that $\log \left( |r - r'|/R_0 \right)$ corresponds to free boundary conditions, where $R_0$ is a constant with dimensions of length.
The vortices move according to the local velocity field,

\[ \frac{d\mathbf{r}_i}{dt} = \mathbf{u}(r_i). \]  

(8)

This could be intuitively clear (the fluid drags the vortices with it). This means

\[ \frac{dx_i}{dt} = \frac{1}{2\pi} \sum_{j \neq i} \Gamma_j (y_i - y_j)/(r_i - r_j)^2 \]  

(9)

\[ \frac{dy_i}{dt} = -\frac{1}{2\pi} \sum_{j \neq i} \Gamma_j (x_i - x_j)/(r_i - r_j)^2. \]

If there is no kinetic energy for the vortices, what are the ‘conjugate variables’ analogous to \(x\) and \(p\) for regular particles?

(d) Check from eqn 9 and eqn 7 that \(\frac{dx_i}{dt} = (1/4\Gamma_i) \frac{\partial H}{\partial y_i}\) and \(\frac{dy_i}{dt} = -(1/4\Gamma_i) \frac{\partial H}{\partial x_i}\). Thus \(2\sqrt{\Gamma_i} x_i\) and \(2\sqrt{\Gamma_i} y_i\) are analogous to \(x\) and \(p\) in a regular phase space.

Note that the continuum vorticity (eqn 6) has total derivative zero \(d\omega/dt = \partial \omega/\partial t + \mathbf{u} \cdot \nabla \omega = 0\).

(d) First, if a packet of fluid travels along a path \(\mathbf{r}(t)\) with \(\dot{\mathbf{r}} = \mathbf{u}(\mathbf{r})\), how will the vorticity \(\omega(\mathbf{r}(t))\) of this packet change with time? Second, in the limit of a \(\delta\)-function vorticity, is this compatible with eqn 8? (Hints: Remember the total derivative is zero.)

Launch the simulation.

(e) Start with \(n = 20\) vortices with a random distribution of vortex strengths \(\Gamma_i \in [-1, 1]\) and random positions \(\mathbf{r}_i\) within the unit circle.\(^5\) Print the original configuration. Run for a time \(t = 10\), and print the final configuration. Do you see the spontaneous formation of a giant whirlpool? Are the final positions roughly also randomly arranged? Measure and report the energy \(H\) of your vortex configuration. (Warning: Sometimes the differential equation solver crashes. Just restart again with a new set of random initial conditions.) Note: we are not keeping the vortices inside the unit circle during our dynamical evolution.

Hurricanes and Jupiter’s Red Spot can be thought of as large concentrations of vorticity – all the plus or minus vortices concentrated into one area, making a giant whirlpool. What do we need to do to arrange this? Let’s consider the effect of the total energy.

(f) For a given set of vortices of strength \(\Gamma_i\), would clustering the positive vortices and the negative vortices each into their own clump (counter-rotating hurricanes) be a high-energy or a low energy configuration, as the size of the clumps goes to zero? (Hint: You will be able to check this with the simulation later.) Why?

\(^5\)You can generate random points in the unit circle by choosing \(\theta \in (-\pi, \pi]\) and picking \(r = \sqrt{p}\) with \(p\) uniformly distributed \(\in [0, 1]\).
You chose points at random for part (e), but did not see the vortices separated into clumps. How did we know this was likely to happen?

(g) *Estimate the entropy difference between a state where twenty vortices are confined within the unit circle, and a state where ten positive vortices and ten negative vortices each are confined inside the particular circles of radius $R = \frac{1}{2}$ shown in Fig. 2.* Leave your answer as a multiple of $k_B$. (Note: The vortices have different values of $\Gamma$, so are distinguishable.) *How many tries would you need to see this clumping happen?*

Onsager’s problem is one of the best examples of *negative temperature*.

(h) *Using your results from part (f) and part (g), in the energy range where the hurricanes will form, will the change in entropy be positive or negative as the energy increases? Is the temperature negative?*

In most simulations of Onsager’s vortices, one selects for states that form a big hurricane by starting with several small ones, and watching them combine. (The small hurricanes must be tightly packed; as they combine they gain entropy because the vortices can spread out more.) Instead, we shall set the energy of our configuration by doing a Metropolis Monte-Carlo simulation at a fixed temperature.

(i) *Thermalize the Monte-Carlo simulation for your $n = 20$ vortices at low temperature $\beta = 1/(k_B T) = 2$, report the final energy, and print out your configuration. Does the thermalized vortex state look similar to the initial conditions you generated in part (e)? Thermalize again at temperature $\beta = -2$, report the energy, and print out your final configuration. Do the vortices separate out into clumps of positive and negative vorticity?*

6It is another interesting question how many small circles could be formed, but that is beyond the scope of this question.

7Because the phase space is finite (footnote ??), the volume of the energy shell at high energies goes to zero, instead of continuing to increase.
The time it takes to run a simulation is roughly determined by the minimum distance between vortices. We can use this to keep our simulation from crashing so much.

(j) Re-run the Monte Carlo simulation with \( n = 20 \) vortices until you find a configuration with a clear separation of positive and negative vortices, but where the minimum distance between vortices is not too small (say, bigger than 0.01). (This should not take lots of tries, so long as you thermalize with an energy in the right region.) Is the energy you need to thermalize to positive, or negative? Print this initial configuration of vortices. Animate the simulation for \( t = 10 \) with this state as the initial condition. How many hurricanes do you find? Print out the final configuration of vortices. (Hint: In Mathematica, one can right-click in the animation window to print the current configuration (Save Graphic As...). In Python, just plot \( xOft[-1], yOft[-1] \); an array evaluated at \([-1]\) gives the last entry. You can copy the axis limits and colors and sizes from the animation, to make a nice plot.)

A.20: Arnol’d cat map, expanded.

1. The Arnol’d cat map, Expanded. (Mathematics)

Why do we suppose equilibrium systems uniformly cover the energy surface? Chaotic motion has sensitive dependence on initial conditions; regions on the energy surface get stretched into thin ribbons that twist and fold in complicated patterns, losing information about the initial conditions and leaving many systems with a uniform distribution of probability over all accessible states. Since energy is the only thing we know is conserved, we average over the energy surface.

Arnol’d developed a simple illustration of this stretching and folding using a function taking a two-dimensional square into itself (called the “cat map”, Figure 3). Liouville’s theorem will tell us that Hamiltonian dynamics preserves the \( 6N \)-dimensional phase-space volume; Arnol’d map preserves area in this \( 1 \times 1 \) square, taking \([0,1) \times [0,1)\) in the plane onto itself:

\[
\Gamma(x, p) = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ p \end{pmatrix} \text{mod } 1 \\
= \begin{pmatrix} 2x + p \\ x + p \end{pmatrix} \text{mod } 1 \\
= \text{Mod}_1 \left( M \begin{pmatrix} x \\ p \end{pmatrix} \right) \tag{10}
\]

where \( M = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \) stretches and squeezes our square ‘energy shell’ and \( \text{Mod}_1 \) cuts and pastes it back back into the square. We imagine a trajectory \((x_{n+1}, p_{n+1}) = \Gamma(x_n, p_n)\) evolving with a discrete ‘time’ \( n \). This map vividly illustrates how Hamiltonian dynamics leads to the microcanonical ensemble.

\footnote{Such maps arise as Poincaré sections of Hamiltonian flows, or as the periodic map given by a Floquet theory. There are many analogies between the behavior of dynamical systems with continuous motion and discrete maps.}
Arnol’d cat map. Arnol’d, in a take-off on Schrödinger’s cat, paints a cat on a 2D phase space, which gets warped and twisted as time evolves. From http://www-chaos.umd.edu/misc/catmap.html.

The cutting and pasting process confuses our analysis as we repeatedly apply the map 
\[(x_n, p_n) = \Gamma^n(x_0, p_0) = \text{Mod}_1(M(\text{Mod}_1(M(\ldots(x_0, p_0))))).\]

Fortunately, we can view the time evolution as many applications of the stretching process, with a final cut-and-paste:
\[(x_n, p_n) = \text{Mod}_1(M^n(x_0, p_0)).\]

(a) Show that the last equation is true. (Hint: Work by induction. You may use the fact that \[y \mod 1 = y + n\] for some integer \(n\) and that \[(z + n) \mod 1 = z \mod 1\] for any integer \(n\).) Thus the stretching by \(M\) can be done in the entire \((x, p)\) plane, and then the resulting thin strip can be folded back into the unit square.

Now we analyze how the map stretches and squeezes the rectangle.

(b) Verify that \((\gamma, 1)\) and \((-1/\gamma, 1)\) are eigenvectors of \(M\), where \(\gamma = (\sqrt{5} + 1)/2\) is the Golden Ratio. What are the eigenvalues? Consider a square box, centered anywhere in the plane, tilted so that its sides are parallel to the two eigenvectors. How does its height and width change under application of \(M\)? Show that its area is conserved.

(Hint: If you are doing this by hand, show that \(1/\gamma = (\sqrt{5} - 1)/2\), which can simplify your algebra.)

The next question explores the map iteration when the initial point is at or near the origin \((x_0, p_0) = (0, 0)\). As illustrated in Figure 4, a small, disk-shaped region of uncertainty, centered on the origin, evolves with time \(n\) into a thin strip. When the iteration would make thin strip hit the square’s boundary, it gets cut in two; further iterations stretch and chop our original circle into a bunch of parallel thin lines. (In the version of our map given by eqn 11, \(M^n\) stretches the initial circle into a growing long thin
Fig. 4 Evolution of an initially concentrated phase-space density. Suppose we know our particle has initial position and momentum confined to a small region \(|(x,p)| < r\). This small region is stretched along an irrational angle under the Arnol’d cat map. (This figure shows the origin \(x = 0, p = 0\) as the center of the figure; in the map of eqn (10) shown in Fig. (3) the origin is at the corner.)

line, which gets folded in by Mod_1 to a series of parallel line segments.)

(c) Calculate the momentum \(h\) at which the thin strip’s midpoint first crosses the line \(x = 0\), and note that \(h\) is irrational. The multiples of an irrational number, taken mod 1, can be shown to be dense in the interval [0,1]. Using this, argue that the resulting thin strip, in the limit of many iterations, will become dense in the unit square (the maximum gap between line segments will go to zero as time goes to infinity).

These lines eventually cover the square (our ‘energy shell’) densely and uniformly – intuitively illustrating how the microcanonical ensemble typically describes the long-term behavior. (The small circle represents an uncertainty in the initial position. As for chaotic motion, Arnol’d’s cat map amplifies the uncertainty in the initial position, which leads to the microcanonical ensemble probability distribution spread uniformly over the square.)

But there are many initial conditions whose trajectories do not cover the energy surface. The most obvious example is the origin, \(x_0 = p_0 = 0\), which maps into itself. In that case, the time average of an operator \(O\), \(\lim_{T \to \infty} 1/T \sum_{n=0}^{T} O(x_n, p_n) = O(0, 0)\), clearly will usually be different from its microcanonical average \(\int_{0}^{1} dx \int_{0}^{1} dp O(x, p)\). The next part of the question asks about a less trivial example of a special initial condition leading to non-ergodic behavior.

(d) If \(x_0\) and \(p_0\) can be written as fractions \(a/q\) and \(b/q\) for positive integers \(q, a \leq q,\)

\(\text{The operator } O \text{ is some function of } x \text{ and } p. \text{ In a physics problem, it could be the energy, the momentum, or the square of the position. In our problem, we’d want } O \text{ to obey the periodic boundary conditions of our square, so it could be something like } O(x, p) = \sin^2(2\pi x)(\cos(4\pi p) + 1).\)
and \( b \leq q \), show that \( x_n \) and \( p_n \) can be written as fractions with denominator \( q \) as well. Show thus that this trajectory must eventually\(^{16}\) settle into a periodic orbit, and give an upper bound to the period. Must the time average of an operator \( O \) for such a trajectory equal the microcanonical average?

You probably have heard that there are infinitely many more irrational numbers than rational ones. So the probability of landing on one of the periodic orbits is zero (points with rational \( x \) and \( p \) coordinates are of measure zero). But perhaps there are even more initial conditions which do not equilibrate?

(e) Find an initial condition with a non-periodic orbit which goes to the origin as \( n \to \infty \). What will the time average of \( O \) be for this initial condition?

Arnol’d’s argument that the cat map \(^{10}\) is ergodic shows that, even taking into account points like those you found in parts (d) and (e), the probability of landing on a non-ergodic component is zero. For ‘almost all’ initial conditions, or any initial condition with a small uncertainty, the Arnol’d cat map trajectory will be dense and uniformly cover its square phase space at long times.

B.1: Random energy model (Extra credit.)

2. Random energy model\(^{11}\) (Disordered systems) \( \triangleright \)

The nightmare of every optimization algorithm is a random landscape; if every new configuration has an energy uncorrelated with the previous ones, no search method is better than systematically examining every configuration. Finding ground states of disordered systems like spin glasses and random-field models, or equilibrating them at non-zero temperatures, is challenging because the energy landscape has many features that are quite random. The random energy model (REM) is a caricature of these disordered systems, where the correlations are completely ignored. While optimization of a single REM becomes hopeless, we shall see that the study of the ensemble of REM problems is quite fruitful and interesting.

The REM has \( M = 2^N \) states for a system with \( N \) ‘particles’ (like an Ising spin glass with \( N \) spins), each state with a randomly chosen energy. It describes systems in limit when the interactions are so strong and complicated that flipping the state of a single particle completely randomizes the energy. The states of the individual particles then need not be distinguished; we label the states of the entire system by \( j \in \{1, \ldots, 2^N\} \).

The energies of these states \( E_j \) are assumed independent, uncorrelated variables with a Gaussian probability distribution

\[
P(E) = \frac{1}{\sqrt{\pi N}} e^{-E^2/N}
\]

of standard deviation \( \sqrt{N/2} \).

\(^{10}\)Since the cat map of eqn (10) is invertible, the orbit will be periodic without an initial transient, but that’s harder to argue.

\(^{11}\)This exercise draws heavily from \([?, \text{chapter 5}]\).
**Microcanonical ensemble.** Consider the states in a small range $E < E_j < E + \delta E$. Let the number of such states in this range be $\Omega(E)\delta E$.

(a) Calculate the average
\[
\langle \Omega(N\epsilon) \rangle_{\text{REM}}
\]
over the ensemble of REM systems, in terms of the energy per particle $\epsilon$. For energies near zero, show that this average density of states grows exponentially as the system size $N$ grows. In contrast, show that $\langle \Omega(N\epsilon) \rangle_{\text{REM}}$ decreases exponentially for $E < -N\epsilon_*$ and for $E > N\epsilon_*$, where the limiting energy per particle
\[
\epsilon_* = \sqrt{\log 2}.
\]

(Hint: The total number of states $2^N$ either grows faster or more slowly than the probability density per state $P(E)$ shrinks.)

What does an exponentially growing number of states mean? Let the entropy per particle be $s(\epsilon) = S(N\epsilon)/N$. Then (setting $k_B = 1$ for notational convenience) $\Omega(E) = \exp(S(E)) = \exp(Ns(\epsilon))$ grows exponentially whenever the entropy per particle is positive.

What does an exponentially decaying number of states for $\epsilon < -\epsilon_*$ mean? It means that, for any particular REM, the likelihood of having any states with energy per particle near $\epsilon$ vanishes rapidly as the number of particles $N$ grows large.

How do we calculate the entropy per particle $s(\epsilon)$ of a typical REM? Can we just use the annealed\(^{12}\) average
\[
s_{\text{annealed}}(\epsilon) = \lim_{N \to \infty} (1/N) \log \langle \Omega(E) \rangle_{\text{REM}}
\]
computed by averaging over the entire ensemble of REMs?

(b) Show that $s_{\text{annealed}}(\epsilon) = \log 2 - \epsilon^2$.

If the energy per particle is above $-\epsilon_*$ (and below $\epsilon_*$), the expected number of states $\Omega(E)\delta E$ grows exponentially with system size, so the fractional fluctuations become unimportant as $N \to \infty$. The typical entropy will become the annealed entropy. On the other hand, if the energy per particle is below $-\epsilon_*$, the number of states in the energy range $(E, E + \delta E)$ rapidly goes to zero, so the typical entropy $s(\epsilon)$ goes to minus infinity. (The annealed entropy is not minus infinity because it gets a contribution from exponentially rare REMs that happen to have an energy level far into the tail of the probability distribution.) Hence
\[
s(\epsilon) = s_{\text{annealed}}(\epsilon) = \log 2 - \epsilon^2 \quad |\epsilon| < \epsilon_*
\]
\[
s(\epsilon) = -\infty \quad |\epsilon| > \epsilon_*.
\]

\(^{12}\)Annealing a disordered system (like an alloy or a disordered metal with frozen-in defects) is done by heating it to allow the defects and disordered regions to reach equilibrium. By averaging $\Omega(E)$ not only over levels within one REM but also over all REMs, we are computing the result of equilibrating over the disorder—an annealed average.
Notice why these arguments are subtle. Each REM model in principle has a different entropy. For large systems as $N \to \infty$, the entropies of different REMs look more and more similar to one another\(^{13}\) (the entropy is self-averaging) whether $|\epsilon| < \epsilon^*$ or $|\epsilon| > \epsilon^*$. However, $\Omega(E)$ is not self-averaging for $|\epsilon| > \epsilon^*$, so the typical entropy is not given by the 'annealed' logarithm $\langle \Omega(E) \rangle_{\text{REM}}$.

This sharp cutoff in the energy distribution leads to a phase transition as a function of temperature.

(c) Plot $s(\epsilon)$ versus $\epsilon$, and illustrate graphically the relation $1/T = \partial S/\partial E = \partial s/\partial \epsilon$ as a tangent line to the curve, using an energy in the range $-\epsilon^* < \epsilon < 0$. What is the critical temperature $T_c$? What happens to the tangent line as the temperature continues to decrease below $T_c$?

When the energy reaches $\epsilon^*$, it stops changing as the temperature continues to decrease (because there are no states\(^{14}\) below $\epsilon^*$).

(d) (Needs some familiarity with parts of chapter 6.) Solve for the free energy per particle $f(T) = \epsilon - Ts$, both in the high-temperature phase and the low temperature phase. (Your formula for $f$ should not depend upon $\epsilon$.) What is the entropy in the low temperature phase? (Warning: The microcanonical entropy is discontinuous at $\epsilon^*$. You’ll need to reason out which limit to take to get the right canonical entropy below $T_c$.)

The REM has a glass transition at $T_c$. Above $T_c$ the entropy is extensive and the REM acts much like an equilibrium system. Below $T_c$ one can show \(^{?}, \text{eqn 5.25}\) that the REM thermal population condenses onto a finite number of states (i.e., a number that does not grow as the size of the system increases), which goes to zero linearly as $T \to 0$.

The mathematical structure of the REM also arises in other, quite different contexts, such as combinatorial optimization (Exercise N.6) and random error correcting codes \(^{?}, \text{chapter 6}\).

\(^{13}\)Mathematically, the entropies per particle of REM models with $N$ particles approach that given by equation \([16]\) with probability one \(^{?}, \text{eqn 5.10}\).

\(^{14}\)The distribution of ground-state energies for the REM is an extremal statistics problem, which for large $N$ has a Gumbel distribution (Exercise N.3).