Unpublished Exercises

Statistical Mechanics: Entropy, Order Parameters, and Complexity

James P. Sethna
Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, NY 14853-2501

CLARENDON PRESS • OXFORD
2006

Contents

List of figures




Unpublished Exercises

N

Exercises

These exercises will likely be included in a later edition of the text, *Statistical Mechanics: Entropy, Order Parameters, and Complexity*, by James P. Sethna (Oxford University Press, http://www.physics.cornell.edu/sethna/StatMech).

(N.1) **The Greenhouse effect and cooling coffee.** ②

Vacuum is an excellent insulator. This is why the surface of the Sun can remain hot ($T_S = 6000^\circ$ K) even though it faces directly onto outer space at the microwave background radiation temperature $T_{MB} = 2.725$ K, (Exercise 7.15). The main way[1] in which heat energy can pass through vacuum is by thermal electromagnetic radiation (photons). We will see in Exercise 7.7 that a black body radiates an energy σT^4 per square meter per second, where $\sigma = 5.67 \times 10^{-8}\,\mathrm{J/(s\,m^2\,K^4)}$.

A vacuum flask or Thermos bottle™ keeps coffee warm by containing the coffee in a *Dewar*—a double-walled glass bottle with vacuum between the two walls.

(a) Coffee at an initial temperature $T_H(0) = 100^\circ$ C of volume $V = 150$ mL is stored in a vacuum flask with surface area $A = 0.1\,\mathrm{m}^2$ in a room of temperature $T_C = 20^\circ$ C. Write down symbolically the differential equation determining how the difference between the coffee temperature and the room temperature $\Delta(t) = T_H(t) - T_C$ decreases with time, assuming the vacuum surfaces of the dewar are black and remain at the current temperatures of the coffee and room. Solve this equation symbolically in the approximation that Δ is small compared to T_c (by approximating $T_H^4 = (T_C + \Delta)^4 \approx T_C^4 + 4\Delta T_C^3$). What is the exponential decay time (the time it take for the coffee to cool by a factor of e*), both symbolically and numerically in seconds?* (Useful conversion: 0° C $= 273.15^\circ$ K.)

Real Dewars are not painted black! They are coated with shiny metals in order to minimize this radiative heat loss. (White or shiny materials not only absorb less radiation, but they also emit less radiation, see exercise 7.7.)

The outward solar energy flux at the Earth's orbit is $\Phi_S = 1370\,\mathrm{W/m^2}$, and the Earth's radius is approximately 6400 km, $r_E = 6.4 \times 10^6$ m. The Earth reflects about 30% of the radiation from the Sun directly back into space (its *albedo* $\alpha \approx 0.3$). The remainder of the energy is eventually turned into heat, and radiated into space again. Like the Sun and the Universe, the Earth is fairly well described as a black-body radiation source in the infrared. We will see in Exercise 7.7 that a black body radiates an energy σT^4 per square meter per second, where $\sigma = 5.67 \times 10^{-8}\,\mathrm{J/(s\,m^2\,K^4)}$.

(b) *What temperature T_A does the Earth radiate at, in order to balance the energy flow from the Sun after direct reflection is accounted for? Is that hotter or colder than you would estimate from the temperatures you've experienced on the Earth's surface?* (Warning: The energy flow in is proportional to the Earth's cross-sectional area, while the energy flow out is proportional to its surface area.)

[1] The sun and stars can also radiate energy by emitting neutrinos. This is particularly important during a supernova.

The reason the Earth is warmer than would be expected from a simple radiative energy balance is the *greenhouse effect*.[2] The Earth's atmosphere is opaque in most of the infrared region in which the Earth's surface radiates heat. (This frequency range coincides with the vibration frequencies of molecules in the Earth's upper atmosphere. Light is absorbed to create vibrations, collisions can exchange vibrational and translational (heat) energy, and the vibrations can later again emit light.) Thus it is the Earth's atmosphere which radiates at the temperature T_A you calculated in part (b); the upper atmosphere has a temperature intermediate between that of the Earth's surface and interstellar space.

The vibrations of oxygen and nitrogen, the main components of the atmosphere, are too symmetric to absorb energy (the transitions have no dipole moment), so the main greenhouse gases are water, carbon dioxide, methane, nitrous oxide, and chlorofluorocarbons (CFCs). The last four have significantly increased due to human activities; CO_2 by $\sim$ 30% (due to burning of fossil fuels and clearing of vegetation), CH_4 by $\sim$ 150% (due to cattle, sheep, rice farming, escape of natural gas, and decomposing garbage), N_2O by $\sim$ 15% (from burning vegetation, industrial emission, and nitrogen fertilizers), and CFCs from an initial value near zero (from former aerosol sprays, now banned to spare the ozone layer). Were it not for the Greenhouse effect, we'd all freeze (like Mars)—but we could overdo it, and become like Venus (whose deep and CO_2-rich atmosphere leads to a surface temperature hot enough to melt lead).

(N.2) **The Dyson sphere.** (Astrophysics) ②

Life on Earth can be viewed as a heat engine, taking energy a hot bath (the Sun at temperature $T_S = 6000^\circ$ K) and depositing it into a cold bath (interstellar space, at a microwave background temperature $T_{MB} = 2.725$ K, Exercise 7.15). The outward solar energy flux at the Earth's orbit is $\Phi_S = 1370\,\mathrm{W/m^2}$, and the Earth's radius is approximately 6400 km, $r_E = 6.4 \times 10^6$ m.

(a) *If life on Earth were perfectly efficient (a Carnot cycle with a hot bath at T_S and a cold bath at T_{MB}), how much useful work (in watts) could be extracted from this energy flow? Compare that to the estimated world marketed energy consumption of* 4.5×10^{20} J/year. (Useful constant: There are about $\pi \times 10^7$ s in a year.)

Your answer to part (a) suggests that we have some ways to go before we run out of solar energy. But let's think big.

(b) *If we built a sphere enclosing the Sun at a radius equal to Earth's orbit (about* 150 *million kilometers, $R_{ES} \approx 1.5 \times 10^{11}$ m), by what factor would the useful work available to our civilization increase?*

This huge construction project is called a *Dyson sphere*, after the physicist who suggested [4] that we look for advanced civilizations by watching for large sources of infrared radiation.

Earth, however, does not radiate at the temperature of interstellar space. It radiates roughly as a black body at near $T_E = 300^\circ\,\mathrm{K} = 23^\circ\,\mathrm{C}$ (see, however, Exercise N.1).

(c) *How much less effective are we at extracting work from the solar flux, if our heat must be radiated effectively to a* 300° K *cold bath instead of one at T_{MB}, assuming in both cases we run Carnot engines?*

There is an alternative point of view, though, which tracks entropy rather than energy. Living beings maintain and multiply their low-entropy states by dumping the entropy generated into the energy stream leading from the Sun to interstellar space. New memory storage also intrinsically involves entropy generation (Exercise 5.2); as we move into the information age, we may eventually care more about dumping entropy than about generating work. In analogy to the 'work effectiveness' of part (c) (ratio of actual work to the Carnot upper bound on the work, given the hot and cold baths), we can estimate an entropy-dumping effectiveness (the ratio of the actual entropy added to the energy stream, compared to the entropy that could be conceivably added given the same hot and cold baths).

(d) *How much entropy impinges on the Earth (per second per square meter) from the Sun? How much leaves the Earth when the solar energy flux is radiated away at temperature* $T_E = 300^\circ$ K*? By what factor f is the entropy dumped*

[2] The glass in greenhouses also is transparent in the visible and opaque in the infrared. This, it turns out, isn't why it gets warm inside; the main insulating effect of the glass is to forbid the warm air from escaping. The greenhouse effect is in that sense poorly named.

to outer space less than the entropy we could dump into a heat bath at T_{MB}? From an entropy-dumping standpoint, which is more important, the hot-bath temperature T_S or the cold-bath temperature (T_E or T_{MB}, respectively)?

For generating useful work, the Sun is the key and the night sky is hardly significant. For dumping the entropy generated by civilization, though, the night sky is the giver of life and the realm of opportunity. These two perspectives are not really at odds. For some purposes, a given amount of work energy is much more useful at low temperatures. Dyson later speculated about how life could make efficient use of this by running at much colder tempeartures (Exercise 5.1). A hyper-advanced information-based civilization would hence want not to radiate in the infrared, but in the microwave range.

To do this, it needs to increase the area of the Dyson sphere; a bigger sphere can re-radiate the Solar energy flow as black-body radiation at a lower temperature. Interstellar space is a good insulator, and one can only shove so much heat energy through it to get to the Universal cold bath. A body at temperature T radiates the largest possible energy if it is completely black. We will see in Exercise 7.7 that a black body radiates an energy σT^4 per square meter per second, where $\sigma = 5.67 \times 10^{-8}\,\mathrm{J}/(\mathrm{s\,m^2\,K^4})$ is the Stefan–Boltzmann constant.

(e) *How large a radius R_D must the Dyson sphere have to achieve 50% entropy-dumping effectiveness? How does this radius compare to the distance to Pluto ($R_{PS} \approx 6 \times 10^{12}\,\mathrm{m}$)? If we measure entropy in bits (using $k_S = (1/\log 2)$ instead of $k_B = 1.3807 \times 10^{-23}\,\mathrm{J/K}$), how many bits per second of entropy can our hyper-advanced civilization dispose of?* (You may ignore the relatively small entropy impinging from the Sun onto the Dyson sphere, and ignore both the energy and the entropy from outer space.)

The sun wouldn't be bright enough to read by at that distance, but if we had a well-insulated sphere we could keep it warm inside—only the outside need be cold. Alternatively, we could just build the sphere for our computers, and live closer in to the Sun; our re-radiated energy would be almost as useful as the original solar energy.

(N.3) **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 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* [10]. 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*

$$(1 - A) \approx \exp(-A) \quad \text{as } N \to \infty, \tag{N.1}$$

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

$$F_N(H) \approx \exp\left[-\exp\left(\frac{\mu - H}{\beta}\right)\right]. \tag{N.2}$$

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

$$\begin{aligned} F(x) &= \exp(-\exp(-x)) \\ \rho(x) &= \partial F/\partial x = \exp(-(x + \exp(-x))). \end{aligned} \tag{N.3}$$

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 [5, section 8.3]. In particular, suppose

$$F_1(H) \approx 1 - A\exp(-BH^\delta) \tag{N.4}$$

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 N.1).

(c) *Show that the extreme value statistics $F_N(H)$ for this distribution is of the Gumbel form (eqn N.2) 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 $F_1(H) = (1/2)(1 + \mathrm{erf}(H/\sqrt{2}))$ 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 N.4 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 [8].

(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 $\rho_N(H) = \mathrm{d}F_N/\mathrm{d}H$ 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 N.8). 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[4] have extreme value distributions that are described by Weibull statistics (see Exercise N.4).

(N.4) **First to fail: Weibull.**[5] (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 N.3 and N.8), 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 $S(t) = \int_0^t \rho(t')\,\mathrm{d}t'$ are often given by the *Weibull distribution*

$$\begin{aligned} S(t) &= \mathrm{e}^{-(t/\alpha)^\gamma}, \\ \rho(t) &= \frac{\mathrm{d}S}{\mathrm{d}t} = -\frac{\gamma}{\alpha}\left(\frac{t}{\alpha}\right)^{\gamma-1}\mathrm{e}^{-(t/\alpha)^\gamma}. \end{aligned} \tag{N.5}$$

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 N.5). 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

$$\rho_1(t) \sim Bt^k \tag{N.6}$$

[3] 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 [10, section 8.2].

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

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

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 N.6) 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 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*,[6] 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.[7] Fracture in brittle materials often occurs due to pre-existing microcracks, typically on the surface of the material. Suppose we have an isolated[8] 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[9] given by

$$\sigma_c(L) \approx K_c/\sqrt{\pi L}. \tag{N.7}$$

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 N.7.)

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 $\rho(L)$?*

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 N.6).

(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 [10, section 8.2], not a Weibull

[6] The Weibull distribution forms a one-parameter family of universality classes; see chapter 12.

[7] 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.

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

[9] 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...

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.

(N.5) **Random energy model.**[10] (Disordered systems) ③

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

$$P(E) = \frac{1}{\sqrt{\pi N}} \mathrm{e}^{-E^2/N} \tag{N.8}$$

of standard deviation $\sqrt{N/2}$.

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*

$$\langle \Omega(N\epsilon) \rangle_{\mathrm{REM}} \tag{N.9}$$

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

$$\epsilon_* = \sqrt{\log 2}. \tag{N.10}$$

(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 in a range near ϵ vanishes rapidly as the energy goes to zero.

How do we calculate the entropy per particle $s(\epsilon)$ of a typical REM? Can we just use the annealed[11] average

$$s_{\mathrm{annealed}}(\epsilon) = \lim_{N \to \infty} (1/N) \log \langle \Omega(E) \rangle_{\mathrm{REM}} \tag{N.11}$$

computed by averaging over the entire ensemble of REMs?

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

If the energy per particle is above $-\epsilon_*$ (and below ϵ_*), the expected number of states $\Omega(E)\,\delta E$ grows exponentially with system size, so the fractional fluctuations become unimportant as

[10] This exercise draws heavily from [5, chapter 5].

[11] 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 equilbrating over the disorder—an annealed average.

$N \to \infty$. 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

$$\begin{aligned} s(\epsilon) &= s_{\text{annealed}}(\epsilon) = \log 2 - \epsilon^2 \quad |\epsilon| < \epsilon_* \\ s(\epsilon) &= -\infty \qquad\qquad\qquad\qquad\quad |\epsilon| > \epsilon_*. \end{aligned} \tag{N.12}$$

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[12] (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 ϵ, 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 happens as the temperature continues to decrease (slope increases)? What is the critical temperature T_c?*

When the energy reaches ϵ_*, it stops changing as the temperature continues to decrease (because there are no states[13] below ϵ_*).

(d) *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 ϵ.) *What is the entropy in the low temperature phase?* (Warning: The microcanonical entropy is discontinuous at ϵ^*. You'll need to reason out which limit to take as $T \to T_c$ from above to get the right canonical entropy.)

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 [5, 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 [5, chapter 6].

(N.6) **A fair split? Number partitioning.**[14] (Computer science, Mathematics, Statistics) ③

A group of N kids want to split up into two teams that are evenly matched. If the skill of each player is measured by an integer, can the kids be split into two groups such that the sum of the skills in each group is the same?

This is the *number partitioning problem* (NPP), a classic and surprisingly difficult problem in computer science. To be specific, it is **NP**–complete—a category of problems for which no known algorithm can guarantee a resolution in a reasonable time (bounded by a polynomial in their size). If the skill a_j of each kid j is in the range $1 \le a_J \le 2^M$, the 'size' of the NPP is defined as NM. Even the best algorithms will, for the hardest instances, take computer time that grows faster than any polynomial in MN, getting exponentially large as the system grows.

In this exercise, we shall explore connections between this numerical problem and the statistical mechanics of disordered systems. Number partitioning has been termed 'the easiest hard problem'. It is genuinely hard numerically; unlike some other **NP**–complete problems, there are no good heuristics for solving NPP (i.e., that work much better than a random search). On the other hand, the random NPP problem (the ensembles of all possible combinations of skills a_j) has many interesting features that can be understood with relatively straightforward arguments and analogies. Parts of the exercise are to be done on the computer; hints can be found on the computer exercises portion of the book Web site [8].

We start with the brute-force numerical approach to solving the problem.

[12] Mathematically, the entropies per particle of REM models with N particles approach that given by equation N.12 with probability one [5, eqn 5.10].

[13] 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).

[14] This exercise draws heavily from [5, chapter 7].

(a) *Write a function* `ExhaustivePartition(S)` *that inputs a list S of N integers, exhaustively searches through the 2^N possible partitions into two subsets, and returns the minimum cost (difference in the sums). Test your routine on the four sets [5] $S_1 = [10, 13, 23, 6, 20]$, $S_2 = [6, 4, 9, 14, 12, 3, 15, 15]$, $S_3 = [93, 58, 141, 209, 179, 48, 225, 228]$, and $S_4 = [2474, 1129, 1388, 3752, 821, 2082, 201, 739]$.*
Hint: S_1 has a balanced partition, and S_4 has a minumum cost of 48. You may wish to return the signs of the minimum-cost partition as part of the debugging process.
What properties emerge from studying ensembles of large partitioning problems? We find a *phase transition*. If the range of integers (M digits in base two) is large and there are relatively few numbers N to rearrange, it is unlikely that a perfect match can be found. (A random instance with $N = 2$ and $M = 10$ has a one chance in $2^{10} = 1024$ of a perfect match, because the second integer needs to be equal to the first.) If M is small and N is large it should be easy to find a match, because there are so many rearrangements possible and the sums are confined to a relatively small number of possible values. It turns out that it is the ratio $\kappa = M/N$ that is the key; for large random systems with $M/N > \kappa_c$ it becomes extremely unlikely that a perfect partition is possible, while if $M/N < \kappa_c$ a fair split is extremely likely.
(b) *Write a function* `MakeRandomPartitionProblem(N,M)` *that generates N integers randomly chosen from $\{1, \ldots, 2^M\}$, rejecting lists whose sum is odd (and hence cannot have perfect partitions). Write a function* `pPerf(N,M,trials)`*, which generates* `trials` *random lists and calls* `ExhaustivePartition` *on each, returning the fraction $p_{\rm perf}$ that can be partitioned evenly (zero cost). Plot $p_{\rm perf}$ versus $\kappa = M/N$, for $N = 3$, 5, 7 and 9, for all integers M with $0 < \kappa = M/N < 2$, using at least a hundred trials for each case. Does it appear that there is a phase transition for large systems where fair partitions go from probable to unlikely? What value of κ_c would you estimate as the critical point?*
Should we be calling this a phase transition? It emerges for large systems; only in the 'thermodynamic limit' where N gets large is the transition sharp. It separates two regions with qualitatively different behavior. The problem is much like a spin glass, with two kinds of random variables: the skill levels of each player a_j are fixed, 'quenched' random variables for a given random instance of the problem, and the assignment to teams can be viewed as spins $s_j = \pm 1$ that can be varied ('annealed' random variables)[15] to minimize the cost $C = |\sum_j a_j s_j|$.
(c) *Show that the square of the cost C^2 is of the same form as the Hamiltonian for a spin glass, $H = \sum_{i,j} J_{ij} s_i s_j$. What is J_{ij}?*
The putative phase transition in the optimization problem (part (b)) is precisely a zero-temperature phase transition for this spin-glass Hamiltonian, separating a phase with zero ground-state energy from one with non-zero energy in the thermodynamic limit.
We can understand both the value κ_c of the phase transition and the form of $p_{\rm perf}(N, M)$ by studying the distribution of possible 'signed' costs $E_{\mathbf{s}} = \sum_j a_j s_j$. These energies are distributed over a maximum total range of $E_{\rm max} - E_{\rm min} = 2 \sum_{j=1}^{2^N} a_j \leq 2N\, 2^M$ (all players playing on the plus team, through all on the minus team). For the bulk of the possible team choices $\{s_j\}$, though, there will be some cancellation in this sum. The probability distribution $P(E)$ of these energies for a particular NPP problem $\{a_j\}$ is not simple, but the average probability distribution $\langle P(E) \rangle$ over the ensemble of NPP problems can be estimated using the central limit theorem. (Remember that the central limit theorem states that the sum of N random variables with mean zero and standard deviation σ converges rapidly to a normal (Gaussian) distribution of standard deviation $\sqrt{N}\sigma$.)
(d) *Estimate the mean and variance of a single term $s_j a_j$ in the sum, averaging over both the spin configurations s_j and the different NPP problem realizations $a_j \in [1, \ldots, 2^M]$, keeping only the most important term for large M.*
(Hint: Approximate the sum as an integral, or use the explicit formula $\sum_1^K k^2 = k^3/3 +$

[15] *Quenched* random variables are fixed terms in the definition of the system, representing dirt or disorder that was frozen in as the system was formed (say, by quenching the hot liquid material into cold water, freezing it into a disordered configuration). *Annealed* random variables are the degrees of freedom that the system can vary to explore different configurations and minimize its energy or free energy.

$k^2/2 + k/6$ and keep only the most important term.) *Using the central limit theorem, what is the ensemble-averaged probability distribution $P(E)$ for a team with N players?* Hint: Here $P(E)$ is non-zero only for *even* integers E, so for large N $P(E) \approx (2/\sqrt{2\pi}\sigma)\exp(-E^2/2\sigma^2)$; the normalization is doubled.

Your answer to part (d) should tell you that the possible energies are mostly distributed among integers in a range of size $\sim 2^M$ around zero, up to a factor that goes as a power of N. The total number of states explored by a given system is 2^N. So, the expected number of zero-energy states should be large if $N \gg M$, and go to zero rapidly if $N \ll M$. Let us make this more precise.

(e) *Assuming that the energies for a specific system are randomly selected from the ensemble average $P(E)$, calculate the expected number of zero-energy states as a function of M and N for large N. What value of $\kappa = M/N$ should form the phase boundary separating likely from unlikely fair partitions? Does that agree well with your numerical estimate from part (b)?*

The assumption we made in part (e) ignores the correlations between the different energies due to the fact that they all share the same step sizes a_j in their random walks. Ignoring these correlations turns out to be a remarkably good approximation.[16] We can use the random-energy approximation to estimate $p_{\rm perf}$ that you plotted in part (b).

(f) *In the random-energy approximation, argue that $p_{\rm perf} = 1 - (1 - P(0))^{2^{N-1}}$. Approximating $(1 - A/L)^L \approx \exp(-A)$ for large L, show that*

$$p_{\rm perf}(\kappa, N) \approx 1 - \exp\left[-\sqrt{\frac{3}{2\pi N}}2^{-N(\kappa-\kappa_c)}\right]. \tag{N.13}$$

Rather than plotting the theory curve through each of your simulations from part (b), we change variables to $x = N(\kappa-\kappa_c)+(1/2)\log_2 N$, where the theory curve

$$p_{\rm perf}^{\rm scaling}(x) = 1 - \exp\left[-\sqrt{\frac{3}{2\pi}}2^{-x}\right] \tag{N.14}$$

is independent of N. If the theory is correct, your curves should converge to $p_{\rm perf}^{\rm scaling}(x)$ as N becomes large

(g) *Reusing your simulations from part (b), make a graph with your values of $p_{\rm perf}(x, N)$ versus x and $p_{\rm perf}^{\rm scaling}(x)$. Does the random-energy approximation explain the data well?*

Rigorous results show that this random-energy approximation gives the correct value of κ_c. The entropy of zero-cost states below κ_c, the probability distribution of minimum costs above κ_c (of the Weibull form, exercise N.4), and the probability distribution of the k lowest cost states are also correctly predicted by the random-energy approximation. It has also been shown that the correlations between the energies of different partitions vanish in the large (N, M) limit so long as the energies are not far into the tails of the distribution, perhaps explaining the successes of ignoring the correlations.

What does this random-energy approximation imply about the computational difficulty of NPP? If the energies of different spin configurations (arrangements of kids on teams) were completely random and independent, there would be no better way of finding zero-energy states (fair partitions) than an exhaustive search of all states. This perhaps explains why the best algorithms for NPP are not much better than the exhaustive search you implemented in part (a); even among **NP**–complete problems, NPP is unusually unyielding to clever methods.[17] It also lends credibility to the conjecture in the computer science community that $\mathbf{P} \neq \mathbf{NP}$–complete; any polynomial-time algorithm for NPP would have to ingeneously make use of the seemingly unimportant correlations between energy levels.

[16] More precisely, we ignore correlations between the energies of different teams $\mathbf{s} = \{s_i\}$, *except* for swapping the two teams $\mathbf{s} \to -\mathbf{s}$. This leads to the $N - 1$ in the exponent of the exponent for $p_{\rm perf}$ in part (f). Notice that in this approximation, NPP is a form of the random energy model (REM, exercise N.5), except that we are interested in states of energy near $E = 0$, rather than minimum energy states.

[17] The computational cost does peak near $\kappa = \kappa_c$. For small $\kappa \ll \kappa_c$ it's relatively easy to find a good solution, but this is mainly because there are so many solutions; even random search only needs to sample until it finds one of them. For $\kappa > \kappa_c$ showing that there is no fair partition becomes slightly easier as κ grows [5, fig 7.3].

(N.7) **Fracture nucleation: elastic theory has zero radius of convergence.**[18] (Condensed matter) ③

In this exercise, we shall use methods from quantum field theory to tie together two topics which American science and engineering students study in their first year of college: Hooke's law and the convergence of infinite series.

Consider a large steel beam, stretched by a moderate strain $\epsilon = \Delta L/L$ (Figure N.1). You may assume $\epsilon \ll 0.1\%$, where we can ignore plastic deformation.

(a) *At non-zero temperature, what is the equilibrium ground state for the beam as $L \to \infty$ for fixed ϵ?* (Hints: Remember, or show, that the free energy per unit (undeformed) volume of the beam is $\frac{1}{2}Y\epsilon^2$. Notice figure N.2 as an alternative candidate for the ground state.) *For steel, with $Y = 2 \times 10^{11}\,\mathrm{N/m^2}$, $\gamma \approx 2.5\,\mathrm{J/m^2}$,*[19] *and density $\rho = 8000\,\mathrm{kg/m^3}$, how much can we stretch a beam of length $L = 10\,m$ before the equilibrium length is broken in two? How does this compare with the amount the beam stretches under a load equal to its own weight?*

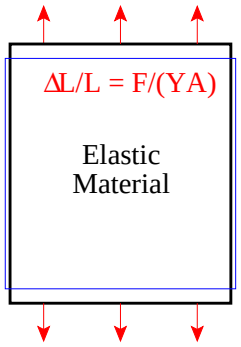


Fig. N.1 Stretched block of elastic material, length L and width W, elongated vertically by a force F per unit area A, with free side boundaries. The block will stretch a distance $\Delta L/L = F/YA$ vertically and shrink by $\Delta W/W = \sigma\,\Delta L/L$ in both horizontal directions, where Y is Young's modulus and σ is Poisson's ratio, linear elastic constants characteristic of the material. For an isotropic material, the other elastic constants can be written in terms of Y and σ; for example, the (linear) bulk modulus $\kappa_{\mathrm{lin}} = Y/3(1-2\sigma)$.

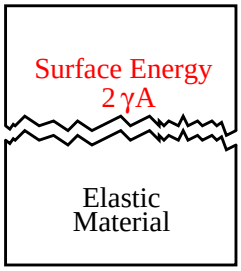
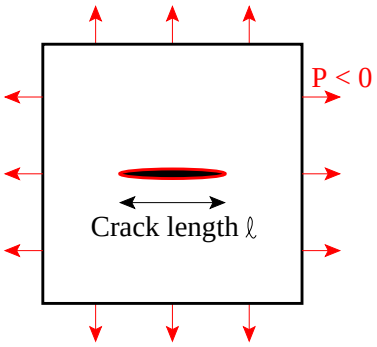


Fig. N.2 Fractured block of elastic material, as in figure N.1 but broken in two. The free energy here is $2\gamma A$, where γ is the free energy per unit area A of (undeformed) fracture surface.

Why don't bridges fall down? The beams in the bridge are in a *metastable state*. What is the barrier separating the stretched and fractured beam states? Consider a crack in the beam, of length ℓ. Your intuition may tell you that tiny cracks will be harmless, but a long crack will tend to grow at small external stress.

For convenient calculations, we will now switch problems from a stretched steel beam to a taut two-dimensional membrane under an isotropic tension, a negative pressure $P < 0$. That is, we are calculating the rate at which a balloon will spontaneously pop due to thermal fluctuations.



[18] This exercise draws heavily on Alex Buchel's work [1, 2].

[19] This is the energy for a clean, flat [100] surface, a bit more than 1eV/surface atom [9]. The surface left by a real fracture in (ductile) steel will be rugged and severely distorted, with a much higher energy per unit area. This is why steel is much harder to break than glass, which breaks in a brittle fashion with much less energy left in the fracture surfaces.

Fig. N.3 Critical crack of length ℓ, in a two-dimensional material under isotropic tension (negative hydrostatic pressure $P < 0$).

The crack costs a surface free energy $2\alpha\ell$, where α is the free energy per unit length of membrane perimeter. A detailed elastic theory calculation shows that a straight crack of length ℓ will release a (Gibbs free) energy $\pi P^2(1-\sigma^2)\ell^2/4Y$.
(b) *What is the critical length ℓ_c of the crack, at which it will spontaneously grow rather than heal? What is the barrier $B(P)$ to crack nucleation? Write the net free energy change in terms of ℓ, ℓ_c, and α. Graph the net free energy change ΔG due to the the crack, versus its length ℓ.*
The point at which the crack is energetically favored to grow is called the *Griffiths threshold*, of considerable importance in the study of brittle fracture.
The predicted fracture nucleation rate $R(P)$ per unit volume from homogeneous thermal nucleation of cracks is thus

$$R(P) = (\text{prefactors}) \exp(-B(P)/k_B T). \quad \text{(N.15)}$$

One should note that thermal nucleation of fracture in an otherwise undamaged, undisordered material will rarely be the dominant failure mode. The surface tension is of order an eV per bond ($> 10^3$ °K/Å), so thermal cracks of area larger than tens of bond lengths will have insurmountable barriers even at the melting point. Corrosion, flaws, and fatigue will ordinarily lead to structural failures long before thermal nucleation will arise.

Advanced topic: Elastic theory has zero radius of convergence.
Many perturbative expansions in physics have zero radius of convergence. The most precisely calculated quantity in physics is the gyromagnetic ratio of the electron [7]

$$\begin{aligned}(g-2)_{\text{theory}} = &\alpha/(2\pi) - 0.328478965\ldots(\alpha/\pi)^2 \\ &+ 1.181241456\ldots(\alpha/\pi)^3 \\ &- 1.4092(384)(\alpha/\pi)^4 \\ &+ 4.396(42)\times 10^{-12} \end{aligned} \quad \text{(N.16)}$$

a power series in the fine structure constant $\alpha = e^2/\hbar c = 1/137.039999\ldots$. (The last term is an α-independent correction due to other kinds of interactions.) Freeman Dyson gave a wonderful argument that this power-series expansion, and quantum electrodynamics as a whole, has zero radius of convergence. He noticed that the theory is sick (unstable) for any negative α (corresponding to a pure imaginary electron charge e). The series must have zero radius of convergence since any circle in the complex plane about $\alpha = 0$ includes part of the sick region.
How does Dyson's argument connect to fracture nucleation? Fracture at $P < 0$ is the kind of instability that Dyson was worried about for quantum electrodynamics for $\alpha < 0$. It has implications for the convergence of nonlinear elastic theory.
Hooke's law tells us that a spring stretches a distance proportional to the force applied: $x - x_0 = F/K$, defining the spring constant $1/K = dx/dF$. Under larger forces, the Hooke's law will have corrections with higher powers of F. We could define a 'nonlinear spring constant' $K(F)$ by

$$\frac{1}{K(F)} = \frac{x(F)-x(0)}{F} = k_0 + k_1 F + \ldots \quad \text{(N.17)}$$

Instead of a spring constant, we'll calculate a nonlinear version of the bulk modulus $\kappa_{\text{nl}}(P)$ giving the pressure needed for a given fractional change in volume, $\Delta P = -\kappa \Delta V/V$. The linear isothermal bulk modulus[20] is given by $1/\kappa_{\text{lin}} = -(1/V)(\partial V/\partial P)|_T$; we can define a nonlinear generalization by

$$\begin{aligned}\frac{1}{\kappa_{\text{nl}}(P)} &= -\frac{1}{V(0)}\frac{V(P)-V(0)}{P} \\ &= c_0 + c_1 P + c_2 P^2 + \cdots + c_N P^N + \cdots \end{aligned} \quad \text{(N.18)}$$

This series can be viewed as higher and higher-order terms in a nonlinear elastic theory.
(c) *Given your argument in part (a) about the stability of materials under tension, would Dyson argue that the series in eqn N.18 has a zero or a non-zero radius of convergence?*
In Exercise 1.5 we saw the same argument holds for Stirling's formula for $N!$, when extended

[20] Warning: For many purposes (e.g. sound waves) one must use the *adiabatic* elastic constant $1/\kappa = -(1/V)(\partial V/\partial P)|_S$. For most solids and liquids these are nearly the same.

to a series in $1/N$; any circle in the complex $1/N$ plane contains points $1/(-N)$ from large negative integers, where we can show that $(-N)! = \infty$. These series are *asymptotic expansions*. Convergent expansions $\sum c_n x^n$ converge for fixed x as $n \to \infty$; asymptotic expansions need only converge to order $O(x^{n+1})$ as $x \to 0$ for fixed n. Hooke's law, Stirling's formula, and quantum electrodynamics are examples of how important, powerful, and useful asymptotic expansions can be.

Buchel [1, 2], using a clever trick from field theory [12, Chapter 40], was able to calculate the large-order terms in elastic theory, essentially by doing a Kramers–Krönig transformation on your formula for the decay rate (eqn N.15) in part (b). His logic works as follows.

• The Gibbs free energy density $\mathcal{G}$ of the metastable state is complex for negative P. The real and imaginary parts of the free energy for complex P form an analytic function (at least in our calculation) except along the negative P axis, where there is a branch cut.

• Our isothermal bulk modulus for $P > 0$ can be computed in terms of $\mathcal{G} = G/V(0)$. Since $\mathrm{d}G = -S\,\mathrm{d}T + V\,\mathrm{d}P + \mu\,\mathrm{d}N$, $V(P) = (\partial G/\partial P)|_T$ and hence[21]

$$\begin{aligned}\frac{1}{\kappa_{\mathrm{nl}}(P)} &= -\frac{1}{V(0)}\frac{(\partial G/\partial P)|_T - V(0)}{P} \\ &= -\frac{1}{P}\left(\left.\frac{\partial \mathcal{G}}{\partial P}\right|_T - 1\right). \end{aligned} \tag{N.19}$$

(d) *Write the coefficients c_n of eqn N.18 in terms of the coefficients g_m in the nonlinear expansion*

$$\mathcal{G}(P) = \sum g_m P^m. \tag{N.20}$$

• The decay rate $R(P)$ per unit volume is proportional to the imaginary part of the free energy $\mathrm{Im}[\mathcal{G}(P)]$, just as the decay rate Γ for a quantum state is related to the imaginary part $i\hbar\Gamma$ of the energy of the resonance. More specifically, for $P < 0$ the imaginary part of the free energy jumps as one crosses the real axis:

$$\mathrm{Im}[\mathcal{G}(P \pm \imath\epsilon)] = \pm(\text{prefactors})R(P). \tag{N.21}$$

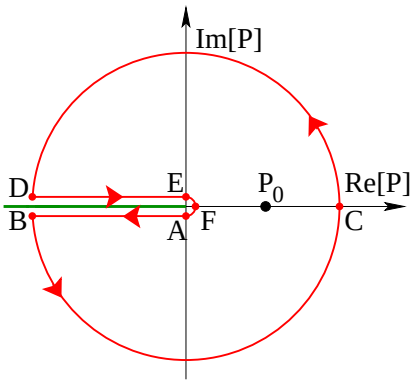


Fig. N.4 Contour integral in complex pressure. The free energy density $\mathcal{G}$ of the elastic membrane is analytic in the complex P plane except along the negative P axis. This allows one to evaluate $\mathcal{G}$ at positive pressure P_0 (where the membrane is stable and $\mathcal{G}$ is real) with a contour integral as shown.

• Buchel then used Cauchy's formula to evaluate the real part of $\mathcal{G}$ in terms of the imaginary part, and hence the decay rate R per unit volume:

$$\begin{aligned}\mathcal{G}(P_0) =& \frac{1}{2\pi\imath}\oint_{ABCDEF}\frac{\mathcal{G}(P)}{P-P_0}\,\mathrm{d}P \\ =& \frac{1}{2\pi\imath}\int_B^0 \frac{\mathcal{G}(P+\imath\epsilon) - \mathcal{G}(P-\imath\epsilon)}{P-P_0}\,\mathrm{d}P \\ &+ \int_{EFA} + \int_{BCD} \\ =& \frac{1}{\pi}\int_B^0 \frac{\mathrm{Im}[\mathcal{G}(P+\imath\epsilon)]}{P-P_0}\,\mathrm{d}P \\ &+ (\text{unimportant}) \end{aligned} \tag{N.22}$$

where the integral over the small semicircle vanishes as its radius $\epsilon \to 0$ and the integral over the large circle is convergent and hence unimportant to high-order terms in perturbation theory.

The decay rate (eqn N.15) for $P < 0$ should be of the form

$$R(P) \propto (\text{prefactors})\,\exp(-D/P^2), \tag{N.23}$$

where D is some constant characteristic of the material. (You may use this to check your answer to part (b).)

[21] Notice that this is not the (more standard) pressure-dependent linear bulk modulus, $\kappa_{\mathrm{lin}}(P)$ which is given by $1/\kappa_{\mathrm{lin}}(P) = -(1/V)(\partial V/\partial P)|_T = -(1/V)(\partial^2\mathcal{G}/\partial P^2)|_T$. This would also have a Taylor series in P with zero radius of convergence at $P = 0$, but it has a different interpretation; $\kappa_{\mathrm{nl}}(P)$ is the nonlinear response at $P = 0$, while $\kappa_{\mathrm{lin}}(P)$ is the pressure-dependent linear response.

(e) *Using eqns. N.21, N.22, and N.23, and assuming the prefactors combine into a constant A, write the free energy for $P_0 > 0$ as an integral involving the decay rate over $-\infty < P < 0$. Expanding $1/(P - P_0)$ in a Taylor series in powers of P_0, and assuming one may exchange sums and integration, find and evaluate the integral for g_m in terms of D and m. Calculate from g_m the coefficients c_n, and then use the ratio test to calculate the radius of convergence of the expansion for $1/\kappa_{\rm nl}(P)$, eqn N.18.* (Hints: Use a table of integrals, a computer algebra package, or change variable $P = -\sqrt{D/t}$ to make your integral into the Γ function,

$$\Gamma(z) = (z-1)! = \int_0^\infty t^{z-1} \exp(-t) {\rm d}t. \quad \text{(N.24)}$$

If you wish, you may use the ratio test on every second term, so the radius of convergence is the value $\lim_{n\to\infty} \sqrt{|c_n/c_{n+2}|}$.)

(Why is this approximate calculation trustworthy? Your formula for the decay rate is valid only up to prefactors that may depend on the pressure; this dependence (some power of P) won't change the asymptotic ratio of terms c_n. Your formula for the decay rate is an approximation, but one which becomes better and better for smaller values of P; the integral for the high-order terms g_m (and hance c_n) is concentrated at small P, so your approximation is asymptotically correct for the high order terms.)

Thus the decay rate of the metastable state can be used to calculate the high-order terms in perturbation theory in the stable phase! This is a general phenomena in theories of metastable states, both in statistical mechanics and in quantum physics.

(N.8) **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 N.3) and in the failure of large systems (where the weakest component or flaw leads to failure, see Exercise N.4). 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 $F_1(x) = \int_{-\infty}^{x} \rho(x')\,{\rm d}x'$. Let us define $\rho_N(x)$ to be the probability density that x is the largest of N random variables drawn from the distribution ρ_1, and 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

$$T[g](x) = 2g(ax + b). \quad \text{(N.25)}$$

(This renormalization group is the same as that we use for sums of random variables in Exercise 12.11 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 N.3) is of the form

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

The Weibull distribution (Exercise N.4) is of the form

$$F_{\text{weibull}}(x) = \begin{cases} \exp(-(-x)^\alpha) & x < 0 \\ 1 & x \geq 0 \end{cases}$$
$$g_{\text{weibull}}(x) = \begin{cases} (-x)^\alpha & x < 0 \\ 0 & x \geq 0, \end{cases} \tag{N.26}$$

and the Fréchet distribution is of the form

$$F_{\text{fréchet}}(x) = \begin{cases} 0 & x \leq 0 \\ \exp(-x^{-\alpha}) & x > 0 \end{cases}$$
$$g_{\text{fréchet}}(x) = \begin{cases} \infty & x < 0 \\ x^{-\alpha} & x \geq 0, \end{cases} \tag{N.27}$$

where $\alpha > 0$ in each case.

(b) *Show that these distributions are fixed points for our renormalization-group transformation eqn N.25. What are a and b for each distribution, in terms of α?*

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*

$$g^*(\mu \pm z) = z^{\alpha'} p_\pm(\gamma \log z) \tag{N.28}$$

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 and $p_\pm$ 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*

$$g^*(x) = \mathrm{e}^{-x/\beta} p(x/\gamma) \tag{N.29}$$

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 N.4.) If the probability distribution decays faster than any polynomial (say, exponentially) then the extreme value statistics will be of the Gumbel form [10, section 8.2]. (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 [10]).

(N.9) **Cardiac dynamics.**[22] (Computation, Biology, Complexity) ④

Reading: References [6, 11], Niels Otani, various web pages on cardiac dynamics, http://otani.vet.cornell.edu, and Arthur T. Winfree, 'Varieties of spiral wave behavior: An experimentalist's approach to the theory of excitable media', *Chaos*, **1**, 303-334 (1991). See also spiral waves in Dictyostelium by Bodenschatz and Franck, http://newt.ccmr.cornell.edu/Dicty/diEp47A.mov and http://newt.ccmr.cornell.edu/Dicty/diEp47A.avi.

The cardiac muscle is an excitable medium. In each heartbeat, a wave of excitation passes through the heart, compressing first the atria which pushes blood into the ventricles, and then compressing the ventricles pushing blood into the body. In this exercise we will study simplified models of heart tissue, that exhibit *spiral waves* similar to those found in arrhythmias.

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

An excitable medium is one which, when triggered from a resting state by a small stimulus, responds with a large pulse. After the pulse there is a refractory period during which it is difficult to excite a new pulse, followed by a return to the resting state. The FitzHugh-Nagumo equations provide a simplified model for the excitable heart tissue:[23]

$$\begin{aligned}\frac{\partial V}{\partial t} &= \nabla^2 V + \frac{1}{\epsilon}(V - V^3/3 - W) \\ \frac{\partial W}{\partial t} &= \epsilon(V - \gamma W + \beta),\end{aligned} \tag{N.30}$$

where V is the transmembrane potential, W is the recovery variable, and $\epsilon = 0.2$, $\gamma = 0.8$, and $\beta = 0.7$ are parameters. Let us first explore the behavior of these equations ignoring the spatial dependence (dropping the $\nabla^2 V$ term, appropriate for a small piece of tissue). The dynamics can be visualized in the (V, W) plane.

(a) *Find and plot the* nullclines *of the FitzHugh-Nagumo equations: the curves along which* $\mathrm{d}V/\mathrm{d}t$ *and* $\mathrm{d}W/\mathrm{d}t$ *are zero (ignoring* $\nabla^2 V$*). The intersection of these two nullclines represents the resting state* (V^*, W^*) *of the heart tissue. We apply a stimulus to our model by shifting the transmembrane potential to a larger value—running from initial conditions* $(V^* + \Delta, W^*)$*. Simulate the equations for stimuli* Δ *of various sizes; plot* V *and* W *as a function of time* t*, and also plot* $V(t)$ *versus* $W(t)$ *along with the nullclines. How big a stimulus do you need in order to get a pulse?*

Excitable systems are often close to regimes where they develop spontaneous oscillations. Indeed, the FitzHugh-Nagumo equations are equivalent to the van der Pol equation (which arose in the study of vacuum tubes), a standard system for studying periodic motion.

(b) *Try changing to* $\beta = 0.4$*. Does the system oscillate?* The threshold where the resting state becomes unstable is given when the nullcline intersection lies at the minimum of the V nullcline, at $\beta_c = 7/15$.

Each portion of the tissue during a contraction wave down the heart is stimulated by its neighbors to one side, and its pulse stimulates the neighbor to the other side. This triggering in our model is induced by the Laplacian term $\nabla^2 V$. We simulate the heart on a two-dimensional grid $V(x_i, y_j, t)$, $W(x_i, y_j, t)$, and calculate an approximate Laplacian by taking differences between the local value of V and values at neighboring points.

There are two natural choices for this Laplacian. The five-point discrete Laplacian is generalization of the one-dimensional second derivative, $\partial^2 V/\partial x^2 \approx (V(x+\mathrm{d}x) - 2V(x) + V(x-\mathrm{d}x))/\mathrm{d}x^2$:

$$\begin{aligned}\nabla^2_{[5]} V(x_i, y_i) \approx\ & (V(x_i, y_{i+1}) + V(x_i, y_{i-1}) \\ & + V(x_{i+1}, y_i) + V(x_{i-1}, y_i) \\ & - 4V(x_i, y_i))/\mathrm{d}x^2 \\ \leftrightarrow\ & \frac{1}{\mathrm{d}x^2}\begin{pmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{pmatrix}\end{aligned} \tag{N.31}$$

where $\mathrm{d}x = x_{i+1} - x_i = y_{i+1} - y_i$ is the spacing between grid points and the last expression is the *stencil* by which you multiply the point and its neighbors by to calculate the Laplacian. The nine-point discrete Laplacian has been fine-tuned for improved circularly symmetry, with stencil

$$\nabla^2_{[9]} V(x_i, y_i) \leftrightarrow \frac{1}{\mathrm{d}x^2}\begin{pmatrix} 1/6 & 2/3 & 1/6 \\ 2/3 & -10/3 & 2/3 \\ 1/6 & 2/3 & 1/6 \end{pmatrix}. \tag{N.32}$$

We will simulate our partial-differential equation (PDE) on a square 100×100 grid with a grid spacing $\mathrm{d}x = 1$.[24] As is often done in PDEs, we will use the crude Euler time-step scheme $V(t + \Delta) \approx V(t) + \Delta \partial V/\partial t$ (see Exercise 3.12): we find $\Delta \approx 0.1$ is the largest time step we can get away with. We will use 'no-flow' boundary conditions, which we implement by setting the Laplacian terms on the boundary to zero (the boundaries, uncoupled from the rest of the system, will quickly turn to their resting state). If you are not supplied with example code that does the two-dimensional plots, you may find them at the text web site [8].

(c) *Solve eqn N.30 for an initial condition equal to the fixed-point* (V^*, W^*) *except for a* 10×10 *square at the origin, in which you should apply a stimulus* $\Delta = 3.0$. (Hint: Your simulation should show a pulse moving outward from the origin, disappearing as it hits the walls.)

[23] Nerve tissue is also an excitable medium, modeled using different *Hodgkin-Huxley* equations.

[24] Smaller grids would lead to less grainy waves, but slow down the simulation a lot.

If you like, you can mimic the effects of the sinoatrial (SA) node (your heart's natural pacemaker) by stimulating your heart model periodically (say, with the same 10×10 square). Realistically, your period should be long enough that the old beat finishes before the new one starts. We can use this simulation to illustrate general properties of solving PDEs.

(d) **Accuracy.** *Compare the five and nine-point Laplacians. Does the latter give better circular symmetry?* **Stability.** *After running for a while, double the time step Δ. How does the system go unstable? Repeat this process, reducing Δ until just before it goes nuts. Do you see inaccuracies in the simulation that foreshadow the instability?*

This checkerboard instability is typical of PDEs with too high a time step. The maximum time step in this system will go as $\mathrm{d}x^2$, the lattice spacing squared—thus to make $\mathrm{d}x$ smaller by a factor of two and simulate the same area, you need four times as many grid points and four times as many time points—giving us a good reason for making $\mathrm{d}x$ as large as possible (correcting for grid artifacts by using improved Laplacians). Similar but much more sophisticated tricks have been used recently to spectacularly increase the performance of lattice simulations of the interactions between quarks [3].

As mentioned above, heart arrhythmias are due to spiral waves. To generate spiral waves we need to be able to start up more asymmetric states—stimulating several rectangles at different times. Also, when we generate the spirals, we would like to emulate electroshock therapy by applying a stimulus to a large region of the heart. We can do both by writing code to interactively stimulate a whole rectangle at one time. Again, the code you have obtained from us should have hints for how to do this.

(e) *Add the code for interactively stimulating a general rectangle with an increment to V of size $\Delta = 3$. Play with generating rectangles in different places while other pulses are going by: make some spiral waves. Clear the spirals by giving a stimulus that spans the system.*

There are several possible extensions of this model, several of which involve giving our model spatial structure that mimics the structure of the heart. (One can introduce regions of inactive 'dead' tissue. One can introduce the atrium and ventricle compartments to the heart, with the SA node in the atrium and an AV node connecting the two chambers ...) Niels Otani has an exercise with further explorations of a number of these extensions, which we link to from the Cardiac Dynamics web site.

References

[1] Buchel, A. and Sethna, J. P. (1996). Elastic theory has zero radius of convergence. *Physical Review Letters*, **77**, 1520. 18, 20

[2] Buchel, A. and Sethna, J. P. (1997). Statistical mechanics of cracks: Thermodynamic limit, fluctuations, breakdown, and asymptotics of elastic theory. *Physical Review E*, **55**, 7669. 18, 20

[3] Davies, C. T. H., Follana, E., Gray, A., Lepage, G. P., and *et al.* (2004). High-precision lattice qcd confronts experiment. *Physical Review Letters*, **92**, 022001. 24

[4] Dyson, Freeman (1960). Search for artificial stellar sources of infrared radiation. *Science*, **131**, 1667–1668. N.2

[5] Mézard, M. and Montanari, A. (To be published (2006)). *Constraint satisfaction networks in Physics and Computation*. Clarendon Press, Oxford. N.3, 10, 13, 12, 14, N.6, 17

[6] Otani, N. (2004). Home page. http://otani.vet.cornell.edu. N.9

[7] Sethna, J. P. (1997). Quantum electrodynamics has zero radius of convergence. http://www.lassp.cornell.edu/sethna/Cracks/QED.html. N.7

[8] Sethna, J. P. and Myers, C. R. (2004). *Entropy, Order Parameters, and Complexity* computer exercises: Hints and software. http://www.physics.cornell.edu/sethna/StatMech/ComputerExercises.html. 3, N.6, 24

[9] Skriver, H. (2004). The hls metals database. http://http://databases.fysik.dtu.dk/hlsPT/. 19

[10] Smith, R. L. (To be published (2006)). *Environmental Statistics*. N.3, 3, 9, N.8

[11] Winfree, A. T. (1991). Varieties of spiral wave behavior: An experimentalist's approach to the theory of excitable media. *Chaos*, **1**, 303–34. N.9

[12] Zinn-Justin, J. (1996). *Quantum field theory and critical phenomena (3rd edition)*. Oxford University Press. 20