

Unpublished Exercises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

The author provides this version of this manuscript with the primary intention of making the text accessible electronically—through web searches and for browsing and study on computers. Oxford University Press retains ownership of the copyright. Hard-copy printing, in particular, is subject to the same copyright rules as they would be for a printed book.

CLARENDON PRESS • OXFORD

2011

Copyright James P. Sethna 2011

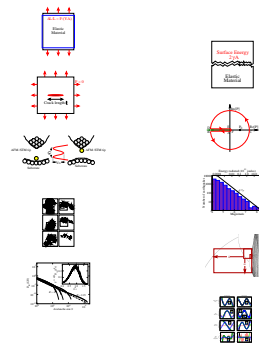
Contents

Contents	v
List of figures	vii
N Unpublished Exercises	1
Exercises	1
N.1 The Greenhouse effect and cooling coffee	1
N.2 The Dyson sphere	2
N.3 Biggest of bunch: Gumbel	3
N.4 First to fail: Weibull	4
N.5 Random energy model	6
N.6 A fair split? Number partitioning	7
N.7 Fracture nucleation: elastic theory has zero radius of convergence	10
N.8 Extreme value statistics: Gumbel, Weibull, and Fréchet	13
N.9 Cardiac dynamics	14
N.10 Quantum dissipation from phonons	16
N.11 The Gutenberg Richter law	17
N.12 Random Walks	17
N.13 Period Doubling	18
N.14 Hysteresis and Barkhausen Noise	19
N.15 Variational mean-field derivation	20
N.16 Avalanche Size Distribution	20
N.17 Ising Lower Critical Dimension	21
N.18 XY Lower Critical Dimension and the Mermin- Wagner Theorem	22
N.19 Long-range Ising	22
N.20 Equilibrium Crystal Shapes	22
N.21 Condition Number and Accuracy	24
N.22 Sherman–Morrison formula	24
N.23 Methods of interpolation	25
N.24 Numerical definite integrals	25
N.25 Numerical derivatives	27
N.26 Summing series	27
N.27 Random histograms	27
N.28 Monte Carlo integration	27
N.29 The Birthday Problem	28
N.30 Washboard Potential	29

N.31	Sloppy Minimization	29
N.32	Sloppy Monomials	30
N.33	Conservative differential equations: Accuracy and fidelity	31
	References	35

List of figures

N.1	Stretched block	10
N.2	Fractured block	10
N.3	Critical crack	11
N.4	Contour integral in complex pressure	12
N.5	Atomic tunneling from a tip	17
N.6	Gutenberg Richter Law	17
N.7	Random walk scaling	18
N.8	Scaling in the period doubling bifurcation diagram	18
N.9	Avalanche size distribution	19
N.10	Interpolation methods	26



Unpublished Exercises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.** (Astrophysics, Ecology) ②

Vacuum is an excellent insulator. This is why the surface of the Sun can remain hot ($T_S = 6000^\circ\text{K}$) even though it faces directly onto outer space at the microwave background radiation temperature $T_{MB} = 2.725\text{K}$, (Exercise 7.15). The main way¹ 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 $\sigma = 5.67 \times 10^{-8}\text{J}/(\text{s m}^2\text{K}^4)$.

A vacuum flask or Thermos bottleTM 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 $T_H(0) = 100^\circ\text{C}$ of volume $V = 150\text{mL}$ is stored in a vacuum flask with surface area $A = 0.1\text{m}^2$ in a room of temperature $T_C = 20^\circ\text{C}$. 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 takes for the coffee to cool by a factor of e), both symbolically and numerically in seconds? (Useful conversion: $0^\circ\text{C} = 273.15^\circ\text{K}$.)

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 $\Phi_S = 1370\text{W}/\text{m}^2$, and the Earth's radius is approximately 6400km , $r_E = 6.4 \times 10^6\text{m}$. 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 $\sigma = 5.67 \times 10^{-8}\text{J}/(\text{s m}^2\text{K}^4)$.

(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.)*

¹The 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; CO₂ by $\sim 30\%$ (due to burning of fossil fuels and clearing of vegetation), CH₄ by $\sim 150\%$ (due to cattle, sheep, rice farming, escape of natural gas, and decomposing garbage), N₂O 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 CO₂-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 $T_S = 6000^\circ\text{K}$) and depositing it into a cold bath (interstellar space, at a microwave background temperature $T_{MB} = 2.725\text{K}$, Exercise 7.15). The outward solar energy flux at the Earth's orbit is $\Phi_S = 1370\text{W/m}^2$, and the Earth's radius is approximately 6400 km, $r_E = 6.4 \times 10^6\text{m}$.

(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 $4.5 \times 10^{20}\text{J/year}$. (Useful constant: There are about $\pi \times 10^7\text{s}$ 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, $R_{ES} \approx 1.5 \times 10^{11}\text{m}$), 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 $T_E = 300^\circ\text{K} = 23^\circ\text{C}$ (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 from the Sun, per second per square meter cross-sectional area? How much leaves the Earth, per second per cross-sectional square meter, when the solar energy flux is radiated away at tem-*

²The 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.

perature $T_E = 300^\circ\text{K}$? By what factor f is the entropy dumped 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 temperatures (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 $\sigma = 5.67 \times 10^{-8} \text{ J}/(\text{s m}^2 \text{ K}^4)$ 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 ($R_{PS} \approx 6 \times 10^{12} \text{ m}$)? If we measure entropy in bits (using $k_S = (1/\log 2)$ instead of $k_B = 1.3807 \times 10^{-23} \text{ J/K}$), 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 $F_1(H) = P(H' < H) = \int^H \rho_1(H') dH'$.

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

The distribution of the largest or smallest of N random numbers is described by *extreme value statistics* [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

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

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

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

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

The constants β and μ just shift the scale and zero of the ruler used to measure the variable of interest. Thus, using a suitable ruler, the largest

of many events is given by a Gumbel distribution

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

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

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

as $H \rightarrow \infty$ 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 $F_1(H) = (1/2)(1 + \text{erf}(H/\sqrt{2}))$ which at large H has asymptotic form $F_1(H) \sim 1 - (1/\sqrt{2\pi}H) \exp(-H^2/2)$. This is of the general form of eqn N.4 with $B = 1/2$ 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 $\rho_N(H) = dF_N/dH$ from part (c). (Hint: $H^*(N) \approx 3.09023$ for $N = 1000$; check this if it is convenient.)

Other types of distributions can have extreme value statistics in different universality classes (see Exercise 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 bound⁴ 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 $S(t) = \int_t^\infty \rho(t') dt'$ are often given by the *Weibull distribution*

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

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 $t \rightarrow 0$, 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 probability $S_N(t)$ at time t that all

³The 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].

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

⁵Developed with the assistance of Paul (Wash) Wawrzynek

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

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

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 isolated⁸ 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 roughly⁹ given by

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

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 $\rho(L)$? (Hint: Consider the population in a small range $d\sigma$, and the same population in the corresponding range dL .)

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

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

⁷Many 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.

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

⁹This 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. . .

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 $x \rightarrow \infty$.)

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 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 $j \in \{1, \dots, 2^N\}$. The energies of these states E_j are assumed independent, uncorrelated variables with a Gaussian probability distribution

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

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

$$\langle \Omega(N\epsilon) \rangle_{\text{REM}} \quad (\text{N.9})$$

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 $\langle \Omega(N\epsilon) \rangle_{\text{REM}}$ decreases exponentially for $E < -N\epsilon_*$ and for $E > N\epsilon_*$, where the limiting energy per particle

$$\epsilon_* = \sqrt{\log 2}. \quad (\text{N.10})$$

(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 with energy per particle near ϵ 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 an-

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

¹¹Annealing 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.

nealed¹¹ average

$$s_{\text{annealed}}(\epsilon) = \lim_{N \rightarrow \infty} (1/N) \log \langle \Omega(E) \rangle_{\text{REM}} \quad (\text{N.11})$$

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 ϵ_*), the expected number of states $\Omega(E) \delta E$ grows exponentially with system size, so the fractional fluctuations become unimportant as $N \rightarrow \infty$. 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

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

Notice why these arguments are subtle. Each REM model in principle has a different entropy. For large systems as $N \rightarrow \infty$, the entropies of different REMs look more and more similar to one another¹² (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 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 ϵ_* , it stops changing as the temperature continues to decrease (because there are no states¹³ 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 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 [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 $T \rightarrow 0$.

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 $1 \leq a_j \leq 2^M$, 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

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

¹³The distribution of ground-state energies for the REM is an extremal statistics problem, which for large N has a Gumbel distribution (Exercise N.3).

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

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.

(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 minimum 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 $\{1, \dots, 2^M\}$, 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_{perf} that can be partitioned evenly

(zero cost). Plot p_{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 $p_{\text{perf}}(N, M)$ by studying the distribution of possible ‘signed’ costs $E_s = \sum_j a_j s_j$. These energies are distributed over a maximum total range of $E_{\text{max}} - E_{\text{min}} = 2 \sum_{j=1}^N a_j \leq 2N 2^M$ (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

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.

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 $a_j \in [1, \dots, 2^M]$, keeping only the most important term for large M . (Hint: Approximate the sum as an integral, or use the explicit formula $\sum_1^K k^2 = K^3/3 + K^2/2 + K/6$ 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_{perf} that you plotted in part (b).

(f) In the random-energy approximation, argue that $p_{\text{perf}} = 1 - (1 - P(0))^{2^{N-1}}$. Approximating

$(1 - A/L)^L \approx \exp(-A)$ for large L , show that

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

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

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

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

(g) Reusing your simulations from part (b), make a graph with your values of $p_{\text{perf}}(x, N)$ versus x and $p_{\text{perf}}^{\text{scaling}}(x)$. 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

¹⁶More precisely, we ignore correlations between the energies of different teams $\mathbf{s} = \{s_i\}$, except for swapping the two teams $\mathbf{s} \rightarrow -\mathbf{s}$. This leads to the $N - 1$ in the exponent of the exponent for p_{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.

¹⁷The 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].

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 ingeniously make use of the seemingly unimportant correlations between energy levels.

(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 cube, 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 cube as $L \rightarrow \infty$ for fixed ϵ ? (Hints: Remember, or show, that the free energy per unit (undeformed) volume of the cube is $\frac{1}{2}Y\epsilon^2$. Notice figure N.2 as an alternative candidate for the ground state.) For steel, with $Y = 2 \times 10^{11} \text{ N/m}^2$, $\gamma \approx 2.5 \text{ J/m}^2$,¹⁹ and density $\rho = 8000 \text{ kg/m}^3$, how much can we stretch a beam of length $L = 10 \text{ m}$ 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?*

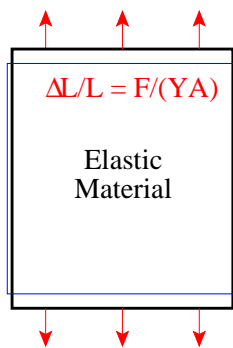


Fig. N.1 Stretched block of elastic material, length L and width W , elongated vertically by a force F per unit area A , with free side boundaries. The block will stretch a distance $\Delta L/L = F/YA$ vertically and shrink by $\Delta W/W = \sigma \Delta L/L$ in both horizontal directions, where Y is 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 $\kappa_{\text{lin}} = Y/3(1 - 2\sigma)$.

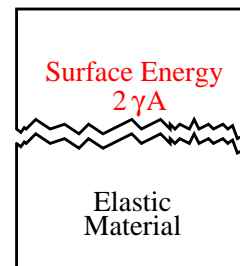


Fig. N.2 Fractured block of elastic material, as in figure N.1 but broken in two. The free energy here is $2\gamma A$, where γ is the free energy per unit area A of (undeformed) fracture surface.

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.

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

¹⁹This 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.

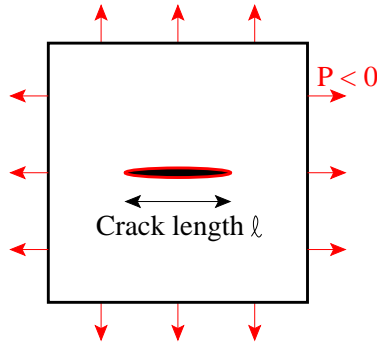


Fig. N.3 Critical crack of length ℓ , in a two-dimensional material under isotropic tension (negative hydrostatic pressure $P < 0$).

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 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, undistorted material will rarely be the dominant failure mode. The surface tension is of order an eV per bond ($> 10^3 \text{ K}/\text{\AA}$), 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]

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

a power series in the fine structure constant $\alpha = e^2/\hbar c = 1/137.035999 \dots$ (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

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

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 modulus²⁰ is given by $1/\kappa_{\text{lin}} = -(1/V)(\partial V/\partial P)|_T$; we can define a nonlinear generalization by

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

²⁰Warning: 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.

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 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 $n \rightarrow \infty$; asymptotic expansions need only converge to order $O(x^{n+1})$ as $x \rightarrow 0$ 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 $dG = -S dT + V dP + \mu dN$, $V(P) = (\partial G / \partial P)|_T$ and hence²¹

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

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

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

- The decay rate $R(P)$ per unit volume is proportional to the imaginary part of the free energy

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

$\text{Im}[\mathcal{G}(P)]$, 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:

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

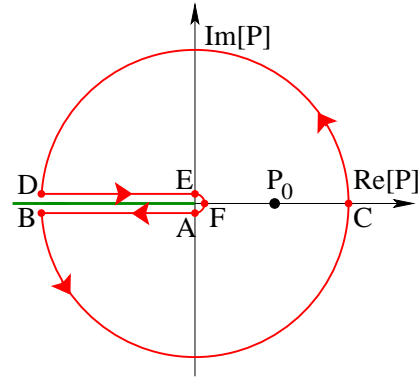


Fig. N.4 Contour integral in complex pressure. The free energy density \mathcal{G} of the elastic membrane is analytic in the complex P plane except along the negative P axis. This allows one to evaluate \mathcal{G} at positive pressure P_0 (where the membrane is stable and \mathcal{G} is real) with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

$$\begin{aligned} \mathcal{G}(P_0) &= \frac{1}{2\pi i} \oint_{ABCDEF} \frac{\mathcal{G}(P)}{P - P_0} dP \\ &= \frac{1}{2\pi i} \int_B^C \frac{\mathcal{G}(P + i\epsilon) - \mathcal{G}(P - i\epsilon)}{P - P_0} dP \\ &\quad + \int_{EFA} + \int_{BCD} \\ &= \frac{1}{\pi} \int_B^C \frac{\text{Im}[\mathcal{G}(P + i\epsilon)]}{P - P_0} dP \\ &\quad + (\text{unimportant}) \end{aligned} \quad (\text{N.22})$$

where the integral over the small semicircle vanishes as its radius $\epsilon \rightarrow 0$ 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

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

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

(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 $1/\kappa_{\text{nl}}(P)$, 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,

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

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

(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 hence 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 $N \rightarrow \infty$ 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 $F_1(x) = \int_{-\infty}^x \rho(x') dx'$. Let us define $\rho_N(x)$ to be the probability density that, out of N random variables, the largest is equal to x . Let $F_N(x)$ to be the corresponding CDF.

(a) Write a formula for $F_{2N}(x)$ in terms of $F_N(x)$. If $F_N(x) = \exp(-g_N(x))$, show that $g_{2N}(x) = 2g_N(x)$.

Our renormalization group coarse-graining operation will remove half of the variables, throwing away the smaller of every pair, and returning the resulting new probability distribution. In terms of the function $g(x) = -\log \int_{-\infty}^x \rho(x') dx'$, it therefore will return a rescaled version of the $2g(x)$. This rescaling is necessary because, as the sample size N increases, the maximum will drift upward—only the form of the probability distribution stays the same, the mean and width can change. Our renormalization-group coarse-graining operation thus maps function space into itself, and is of the form

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

(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

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

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

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

and the Fréchet distribution is of the form

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

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 $x \rightarrow ax + b$ has a fixed point $x = \mu$. Show that the most general form for the fixed-point function is

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

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 , p_+ , and p_- gives the Weibull distribution? The Fréchet distribution?

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

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

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

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 Fréchet 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

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

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.

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

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

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 dV/dt and dW/dt 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 neigh-

bors 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, $\partial^2 V/\partial x^2 \approx (V(x+dx) - 2V(x) + V(x-dx))/dx^2$:

$$\begin{aligned}\nabla_{[5]}^2 V(x_i, y_i) &\approx (V(x_i, y_{i+1}) + V(x_i, y_{i-1}) \\ &\quad + V(x_{i+1}, y_i) + V(x_{i-1}, y_i) \\ &\quad - 4V(x_i, y_i))/dx^2 \\ &\leftrightarrow \frac{1}{dx^2} \begin{pmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{pmatrix}\end{aligned}\quad (\text{N.31})$$

where $dx = x_{i+1} - x_i = y_{i+1} - y_i$ 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

$$\nabla_{[9]}^2 V(x_i, y_i) \leftrightarrow \frac{1}{dx^2} \begin{pmatrix} 1/6 & 2/3 & 1/6 \\ 2/3 & -10/3 & 2/3 \\ 1/6 & 2/3 & 1/6 \end{pmatrix}.\quad (\text{N.32})$$

We will simulate our partial-differential equation (PDE) on a square 100×100 grid with a grid spacing $dx = 1$.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*

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

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

a stimulus $\Delta = 3.0$. (Hint: Your simulation should show a pulse moving outward from the origin, disappearing as it hits the walls.)

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 dx^2 , the lattice spacing squared—thus to make dx 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 dx 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.

(N.10) **Quantum dissipation from phonons.** (Quantum) ②

Electrons cause overlap catastrophes (X-ray edge effects, the Kondo problem, macroscopic quantum tunneling); a quantum transition of a subsystem coupled to an electron bath ordinarily must emit an infinite number of electron-hole excitations because the bath states before and after the transition have zero overlap. This is often called an *infrared catastrophe* (because it is low-energy electrons and holes that cause the zero overlap), or an *orthogonality catastrophe* (even though the two bath states aren’t just orthogonal, they are in different Hilbert spaces). Phonons typically do not produce overlap catastrophes (Debye–Waller, Frank–Condon, Mössbauer). This difference is usually attributed to the fact that there are many more low-energy electron-hole pairs (a constant density of states) than there are low-energy phonons ($\omega_k \sim ck$, where c is the speed of sound and the wave-vector density goes as $(V/2\pi)^3 d^3k$).

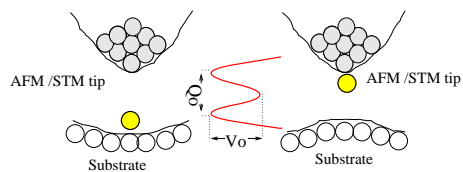


Fig. N.5 Atomic tunneling from a tip. Any *internal* transition among the atoms in an insulator can only exert a force impulse (if it emits a momentum, say into an emitted photon), or a force dipole (if the atomic configuration rearranges); these lead to non-zero phonon overlap integrals only partially suppressing the transition. But a quantum transition that changes the net force between two macroscopic objects (here a surface and a STM tip) can lead to a change in the net force (a force monopole). We ignore here the surface, modeling

the force as exerted directly into the center of an insulating elastic medium.²⁵See “Atomic Tunneling from a STM/AFM Tip: Dissipative Quantum Effects from Phonons” Ard A. Louis and James P. Sethna, *Phys. Rev. Lett.* **74**, 1363 (1995), and “Dissipative tunneling and orthogonality catastrophe in molecular transistors”, S. Braig and K. Flensberg, *Phys. Rev. B* **70**, 085317 (2004).

However, the coupling strength to the low energy phonons has to be considered as well. Consider a small system undergoing a quantum transition which exerts a net force at $x = 0$ onto an insulating crystal:

$$\mathcal{H} = \sum_k p_k^2/2m + 1/2 m\omega_k^2 q_k^2 + F \cdot u_0. \quad (\text{N.33})$$

Let us imagine a kind of scalar elasticity, to avoid dealing with the three phonon branches (two transverse and one longitudinal); we thus naively write the displacement of the atom at lattice site x_n as $u_n = (1/\sqrt{N}) \sum_k q_k \exp(-ikx_n)$ (with N the number of atoms), so $q_k = (1/\sqrt{N}) \sum_n u_n \exp(ikx_n)$.

Substituting for u_0 in the Hamiltonian and completing the square, find the displacement Δ_k of each harmonic oscillator. Write the formula for the likelihood $\langle F|0 \rangle$ that the phonons will all end in their ground states, as a product over k of the phonon overlap integral $\exp(-\Delta_k^2/8a_k^2)$ (with $a_k = \sqrt{\hbar/2m\omega_k}$ the zero-point motion in that mode). Converting the product to the exponential of a sum, and the sum to an integral $\sum_k \sim (V/(2\pi)^3) \int dk$, do we observe an overlap catastrophe?

(N.11) **The Gutenberg Richter law.** (Scaling) ③

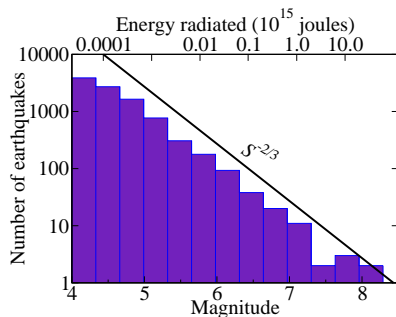


Fig. N.6 Gutenberg Richter Law. The number of earthquakes in 1995 as a function of their magnitude M or their size (energy radiated) S .

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

The size S of an earthquake (the energy radiated, shown in the upper axis of Figure N.6) is traditionally measured in terms of a ‘magnitude’ $M \propto \log S$ (lower axis). The Gutenberg-Richter law tells us that the number of earthquakes of magnitude $M \propto \log S$ goes down as their size S increases. Figure N.6 shows that the number of avalanches of magnitude between M and $M + 1$ is proportional to S^{-B} with $B \approx 2/3$. However, it is traditional in the physics community to consider the probability density $P(S)$ of having an avalanche of size S .

If $P(S) \sim S^{-\tau}$, give a formula for τ in terms of the Gutenberg-Richter exponent B . (Hint: The bins in the histogram have different ranges of size S . Use $P(M) dM = P(S) dS$.)

(N.12) **Random Walks.** (Scaling) ③

Self-similar behavior also emerges without proximity to any obvious transition. One might say that some *phases* naturally have self-similarity and power laws. Mathematicians have a technical term *generic* which roughly translates to ‘without tuning a parameter to a special value’, and so this is termed *generic scale invariance*.

The simplest example of generic scale invariance is that of a random walk. Figure N.7 shows that a random walk appears statistically self-similar.

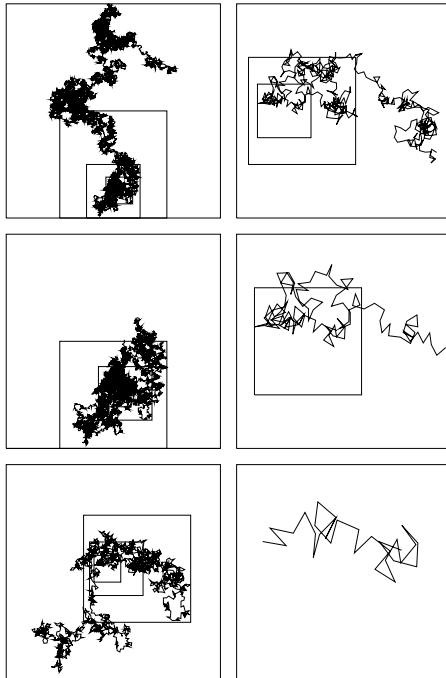


Fig. N.7 Random walk scaling. Each box shows the first quarter of the random walk in the previous box. While each figure looks different in detail, they are *statistically* self-similar. That is, an ensemble of medium-length random walks would be indistinguishable from an ensemble of suitably rescaled long random walks.

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

(N.13) **Period Doubling.** (Scaling) ③

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

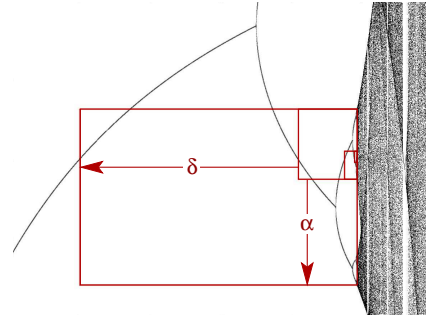


Fig. N.8 Scaling in the period doubling bifurcation diagram. Shown are the points x on the attractor (vertical) as a function of the control parameter μ (horizontal), for the logistic map $f(x) = 4\mu x(1 - x)$, near the transition to chaos.

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

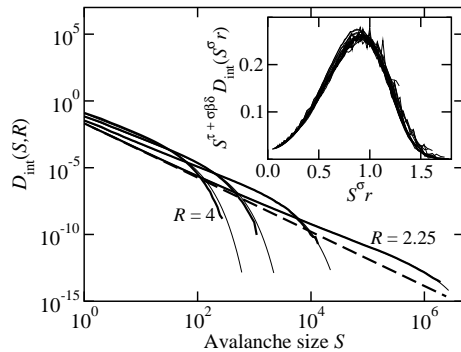
(N.14) **Hysteresis and Barkhausen Noise.** (Scaling) ③

Fig. N.9 Avalanche size distribution, as a function of the disorder, for a model of hysteresis. The thin lines are the prediction of the scaling form eqn (N.34), fit to data near R_c . The inset shows a *scaling collapse*; all the data collapses onto the scaling function $\mathcal{D}(S^\sigma r)$. (The inset uses the notation of the original paper: the probability $D(S, R)$ is called D_{int} because it is ‘integrated’ over magnetic field, τ is called $\tau + \sigma\beta\delta$ and $r = (R_c - R)$).

Hysteresis is associated with abrupt phase transitions. Supercooling and superheating are examples (as temperature crosses T_c). Magnetic recording, the classic place where hysteresis is studied, is also governed by an abrupt phase transition – here the hysteresis in the magnetization, as the external field H is increased (to magnetize the system) and then decreased again to zero. Magnetic hysteresis is characterized by crackling (Barkhausen) electromagnetic noise. This noise is due to avalanches of spins flipping as the magnetic interfaces jerkily are pushed past defects by the external field (much like earthquake faults jerkily responding to the stresses from the tectonic plates). It is interesting that when dirt is added to this abrupt magnetic transition, it exhibits the power-law scaling characteristic of continuous transitions.

Our model of magnetic hysteresis (unlike the experiments) has avalanches and scaling only at a special critical value of the disorder $R_c \sim 2.16$ (Figure N.9). The probability distribution $D(S, R)$ has a power law $D(S, R_c) \propto S^{-\bar{\tau}}$ at the critical point, but away from the critical point takes the *scaling form*

$$D(S, R) \propto S^{-\bar{\tau}} \mathcal{D}(S^\sigma (R - R_c)). \quad (\text{N.34})$$

Note from eqn (N.34) that at the critical disorder $R = R_c$ the distribution of avalanche sizes is a power law $D(S, R_c) = S^{-\bar{\tau}}$. The scaling form controls how this power law is altered as R moves away from the critical point. From Figure N.9 we see that the main effect of moving above R_c is to cut off the largest avalanches at a typical largest size $S_{\text{max}}(R)$, and another important effect is to form a ‘bulge’ of extra avalanches just below the cut-off.

Using the scaling form from eqn N.34, with what exponent does S_{max} diverge as $r = (R_c - R) \rightarrow 0$? (Hint: At what size S is $D(S, R)$, say, one millionth of $S^{-\bar{\tau}}$?) Given $\bar{\tau} \approx 2.03$, how does the mean $\langle S \rangle$ and the mean-square $\langle S^2 \rangle$ avalanche size scale with $r = (R_c - R)$? (Hint: Your integral for the moments should have a lower cutoff S_0 , the smallest possible avalanche, but no upper cutoff, since that is provided by the scaling function \mathcal{D} . Assume $\mathcal{D}(0) > 0$. Change variables to $Y = S^\sigma r$. Which moments diverge?)

Mean Field: Introduction

Mean field theory can be derived and motivated in several ways.

- The interaction (field) from the neighbors can be approximated by the average (mean) field of all sites in the system, effectively as described in Cardy section 2.1. This formulation makes it easy to understand why mean-field theory becomes more accurate in high dimensions and for long-range interactions, as a spin interacts with more and more neighbors.
- The free energy can be bounded above by the free energy of a noninteracting mean-field model. This is based on a variational principle I first learned from Feynman (Exercise N.15, based on Cardy’s exercise 2.1).
- The free energy can be approximated by the contribution of one order parameter configuration, ignoring fluctuations. (From a path-integral point of view, this is a ‘zero loop’ approximation.) For the Ising model, one needs first to change away from the Ising spin variables to some continuous order parameter, either by coarse-graining (as in Ginsburg-Landau theory, Sethna 9.5 below) or by introducing Lagrange multipliers (the Hubbard-Stratonovich transformation, not discussed here).

- (d) The Hamiltonian can be approximated by an infinite-range model, where each spin does interact with the average of all other spins. Instead of an approximate calculation for the exact Hamiltonian, this is an exact calculation for an approximate Hamiltonian – and hence is guaranteed to be at least a sensible physical model. See Exercise N.16 for an application to avalanche statistics.
- (e) The lattice of sites in the Hamiltonian can be approximated as a branching tree (removing the loops) called the *Bethe lattice* (not described here). This yields a different, solvable, mean-field theory, which ordinarily has the same mean-field critical exponents but different non-universal features.

(N.15) **Variational mean-field derivation.** (Mean field) ③

Just as the entropy is maximized by a uniform probability distribution on the energy shell for a fixed-energy (microcanonical) ensemble, it is a known fact that the Boltzmann distribution $\exp(-\beta\mathcal{H})/Z$ minimizes the free energy F for a fixed-temperature (canonical) ensemble described by Hamiltonian energy function \mathcal{H} . If a general ensemble ρ' gives the probability of every state X in a system, and we define for an observable B the average $\langle B \rangle_{\rho'} = \sum_X B(X)\rho'(X)$ then what this says is that

$$\begin{aligned}
 F &= - (1/\beta) \log \sum_X \exp(-\beta\mathcal{H}(X)) \\
 &\leq \langle F \rangle_{\rho'} = \langle E - TS \rangle_{\rho'} \\
 &= \langle \mathcal{H} \rangle_{\rho'} - T \langle S \rangle_{\rho'} \quad (N.35)
 \end{aligned}$$

Feynman, in his work on the polaron problem, derived a variational formula for what appears to be a related problem. If we define $\langle B \rangle_{\mathcal{H}'}$ to be the Boltzmann average of B under the Hamiltonian \mathcal{H}' , Feynman says that

$$\text{Tr} e^{-\beta\mathcal{H}} \geq \text{Tr} e^{-\beta(\mathcal{H}' + \langle \mathcal{H} - \mathcal{H}' \rangle_{\mathcal{H}'})}. \quad (N.36)$$

Feynman’s formula varies not among general probability distributions ρ' , but only among approximating Hamiltonians \mathcal{H}' .

(a) Show that eqn N.35 implies eqn N.36. (Hint: $\langle S \rangle_{\mathcal{H}'}$ is not dependent on \mathcal{H} , and $F' = E' - TS'$.)

(b: Cardy exercise 2.1) Derive the mean field equations for the Ising ferromagnet using Feynman’s inequality (eqn N.36), where \mathcal{H} is the full Hamiltonian and $\mathcal{H}' = h \sum_i S_i$ is a trial Hamiltonian, with h chosen to maximize the right-hand side.

Notice that having a variational bound on the free energy is not as useful as it sounds. The free energy itself is usually not of great interest; the objects of physical importance are almost always derivatives (first derivatives like the pressure, energy and entropy, and second derivatives like the specific heat) or differences (telling which phase is more stable). Knowing an upper bound to the free energy doesn’t tell us much about these other quantities. Like most variational methods, mean-field theory is useful not because of its variational bound, but because (if done well) it provides qualitative insight into the behavior.

(N.16) **Avalanche Size Distribution.** (Scaling function) ③

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

$$\mathcal{H} = -J_0/N \sum_{i,j} S_i S_j - \sum_i (H + h_i) S_i. \quad (N.37)$$

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

$$\begin{aligned}
 H_i^{\text{loc}} &= \frac{\partial \mathcal{H}}{\partial S_i} = J_0/N \sum_j S_j + H + h_i \\
 &= J_0 m + H + h_i
 \end{aligned}$$

changes sign.²⁶ Here $m = (1/N) \sum_j S_j$ is the average magnetization of the system. All spins start by pointing down. A new avalanche is launched when the least stable spin (the unflipped spin of largest local field) is flipped by increasing the external field H . Each spin flip changes the magnetization by $2/N$. If the magnetization change from the first spin flip is enough to trigger the next-least-stable spin, the avalanche will continue.

²⁶We ignore the self-interaction, which is unimportant at large N

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

$$\rho(h) = (1+t)/(2J_0). \quad (\text{N.38})$$

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

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

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

(b) (Optional) By imagining putting periodic boundary conditions on the interval $\{h, h + 2J_0S/N\}$, argue that exactly one spin out of the group of S spins will trigger the rest as a single avalanche. (Hint from Ben Machta: For simplicity, we may assume²⁷ the avalanche starts at $H = m = 0$. Try plotting the local field $H^{\text{loc}}(h') = J_0m(h') + h'$ that a spin with random field h' would feel if the spins between h' and h were flipped. How would this function change if we shuffle all the random fields around the periodic boundary conditions?)

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

$$D(S, t) = \frac{S^{S-1}}{S!} (t+1)^{S-1} e^{-S(t+1)}. \quad (\text{N.39})$$

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

$$D(S, t) = S^{-\tau} \mathcal{D}(S/t^{-x}) \quad (\text{N.40})$$

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

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

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

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

(See "Hysteresis and hierarchies, dynamics of disorder-driven first-order phase transformations", J. P. Sethna, K. Dahmen, S. Kartha, J. A. Krumhansl, B. W. Roberts, and J. D. Shore, PRL **70**, 3347 (1993) for more information.)

(N.17) **Ising Lower Critical Dimension.** (Dimension dependence) ③

What is the lower critical dimension of the Ising model? If the total energy ΔE needed to destroy long-range order is finite as the system size L goes to infinity, and the associated entropy grows with system size, then surely long-range order is possible only at zero temperature.

(a) **Ising model in D dimensions.** Consider the Ising model in dimension D on a hypercubic lattice of length L on each side. Estimate the energy²⁸ needed to create a domain wall splitting the system into two equal regions (one spin up, the other spin down). In what dimension will this wall have finite energy as $L \rightarrow \infty$? Suggest a bound for the lower critical dimension of the Ising model.

The scaling at the lower critical dimension is often unusual, with quantities diverging in ways different from power laws as the critical temperature T_c is approached.

(b) **Correlation length in 1D Ising model.** Estimate the number of domain walls at temperature T in the 1D Ising model. How does the correlation length ξ (the distance between domain walls) grow as $T \rightarrow T_c = 0$? (Hint: Change variables to $\eta_i = S_i S_{i+1}$, which is -1 if there is a domain wall between sites i and $i+1$.) The correlation exponent ν satisfying $\xi \sim (T - T_c)^{-\nu}$ is $1, \sim 0.63, \text{ and } 1/2$ in dimensions $D = 2, 3, \text{ and } \geq 4$, respectively. Is there an exponent ν governing this divergence in one dimension? How does ν behave as $D \rightarrow 1$?

²⁷Equivalently, measure the random fields with respect to h_0 of the triggering spin, and let m be the magnetization change since the avalanche started.

²⁸Energy, not free energy! Think about $T = 0$.

(N.18) **XY Lower Critical Dimension and the Mermin-Wagner Theorem.** (Dimension dependence) ③

Consider a model of continuous unit-length spins (e.g., XY or Heisenberg) on a D -dimensional hypercubic lattice of length L . Assume a nearest-neighbor ferromagnetic bond energy

$$-J \mathbf{S}_i \cdot \mathbf{S}_j. \quad (\text{N.41})$$

Estimate the energy needed to twist the spins at one boundary 180° with respect to the other boundary (the energy difference between periodic and antiperiodic boundary conditions along one axis). In what dimension does this energy stay finite in the thermodynamic limit $L \rightarrow \infty$? Suggest a bound for the lower critical dimension for the emergence of continuous broken symmetries in models of this type.

Note that your argument produces only one thick domain wall (unlike the Ising model, where the domain wall can be placed in a variety of places). If in the lower critical dimension its energy is fixed as $L \rightarrow \infty$ at a value large compared to $k_B T$, one could imagine most of the time the order might maintain itself across the system. The actual behavior of the XY model in its lower critical dimension is subtle.

On the one hand, there cannot be long-range order. This can be seen convincingly, but not rigorously, by estimating the effects of fluctuations at finite temperature on the order parameter, within linear response. Pierre Hohenberg, David Mermin and Herbert Wagner proved it rigorously (including nonlinear effects) using an inequality due to Bogoliubov. One should note, though, that the way this theorem is usually quoted ("continuous symmetries cannot be spontaneously broken at finite temperatures in one and two dimensions") is too general. In particular, for two-dimensional crystals one has long-range order in the crystalline *orientations*, although one does not have long-range broken translational order.

On the other hand, the XY model does have a phase transition in its lower critical dimension at a temperature $T_c > 0$. The high-temperature phase is a traditional paramagnetic phase, with exponentially decaying correlations between orientations as the distance increases. The low-temperature phase indeed lacks long-range order, but it does have a *stiffness* – twisting the system (as in your calculation above) by 180°

costs a free energy that goes to a constant as $L \rightarrow \infty$. In this stiff phase the spin-spin correlations die away not exponentially, but as a power law.

The corresponding *Kosterlitz–Thouless phase transition* has subtle, fascinating scaling properties. Interestingly, the defect that destroys the stiffness (a vortex) in the Kosterlitz–Thouless transition does *not* have finite energy as the system size L gets large. We shall see that its energy grows $\sim \log L$, while its entropy grows $\sim T \log L$, so entropy wins over energy as the temperature rises, even though the latter is infinite.

(N.19) **Long-range Ising.** (Dimension dependence) ③

The one-dimensional Ising model can have a finite-temperature transition if we give each spin an interaction with distant spins.

Long-range forces in the 1d Ising model. Consider an Ising model in one dimension, with long-range ferromagnetic bonds

$$\mathcal{H} = \sum_{i>j} \frac{J}{|i-j|^\sigma} S_i S_j. \quad (\text{N.42})$$

For what values of σ will a domain wall between up and down spins have finite energy? Suggest a bound for the ‘lower critical power law’ for this long-range one-dimensional Ising model, below which a ferromagnetic state is only possible when the temperature is zero. (Hint: Approximate the sum by a double integral. Avoid $i = j$.)

The long-range 1D Ising model at the lower critical power law has a transition that is closely related to the Kosterlitz–Thouless transition. It is in the same universality class as the famous (but obscure) Kondo problem in quantum phase transitions. And it is less complicated to think about and less complicated to calculate with than either of these other two cases.

(N.20) **Equilibrium Crystal Shapes.** (Crystal shapes) ③

What is the equilibrium shape of a crystal? There are nice experiments on single crystals of salt, gold, and lead crystals (see http://www.lassp.cornell.edu/sethna/Crystal_Shapes/Equilibrium_Crystal_Shapes.html).

They show beautiful faceted shapes, formed by carefully annealing single crystals to equilibrium at various temperatures. The physics governing the shape involves the anisotropic surface tension $\gamma(\hat{n})$ of the surface, which depends on the

orientation \hat{n} of the local surface with respect to the crystalline axes.

We can see how this works by considering the problem for atoms on a 2D square lattice with near-neighbor interactions (a lattice gas which we map in the standard way onto a conserved-order parameter Ising model). Here $\gamma(\theta)$ becomes the line tension between the up and down phases – the interfacial free energy per unit length between an up-spin and a down-spin phase. We draw heavily on Craig Rottman and Michael Wortis, *Phys. Rev. B* **24**, 6274 (1981), and on W. K. Burton, N. Cabrera and F. C. Frank, *Phil. Trans. R. Soc. Lond. A* **243** 299-358 (1951).

(a) Interfacial free energy, $T = 0$. Consider an interface at angle θ between up spins and down spins. Show that the energy cost per unit length of the interface is

$$\gamma_0(\theta) = 2J(\cos(|\theta|) + \sin(|\theta|)), \quad (\text{N.43})$$

where length is measured in lattice spacings.

(b) Interfacial free energy, low T . Consider an interface at angle $\theta = \arctan N/M$, connecting the origin to the point (M, N) . At zero temperature, this will be an ensemble of staircases, with N steps upward and M steps forward. Show that the total number of such staircases is $(M + N)!/(M!N!)$. Hint: The number of ways of putting k balls into ℓ jars allowing more than one ball per jar (the ‘number of combinations with repetition’) is $(k + \ell - 1)!/(k!(\ell - 1)!)$. Look up the argument. Using Stirling’s formula, show that the entropy per unit length is

$$s_0(\theta) = k_B \left((\cos(|\theta|) + \sin(|\theta|)) \log(\cos(|\theta|) + \sin(|\theta|)) - \cos(|\theta|) \log(\cos(|\theta|)) - \sin(|\theta|) \log(\sin(|\theta|)) \right). \quad (\text{N.44})$$

How would we generate an equilibrium crystal for the 2D Ising model? (For example, see ‘The Gibbs-Thomson formula at small island sizes

– corrections for high vapour densities” Badrinarayan Krishnamachari, James McLean, Barbara Cooper, and James P. Sethna, *Phys. Rev. B* **54**, 8899 (1996).) Clearly we want a conserved order parameter simulation (otherwise the up-spin ‘crystal’ cluster in a down-spin ‘vapor’ environment would just flip over). The tricky part is that an up-spin cluster in an infinite sea of down-spins will evaporate – it’s unstable.²⁹ The key is to do a simulation below T_c , but with a net (conserved) negative magnetization slightly closer to zero than expected in equilibrium. The extra up-spins will (in equilibrium) mostly find one another, forming a cluster whose time-average will give an equilibrium crystal.

Rottman and Wortis tell us that the equilibrium crystal shape (minimizing the perimeter energy for fixed crystal area) can be found as a parametric curve

$$\begin{aligned} x &= \cos(\theta)\gamma(\theta, T) - \sin(\theta)\frac{d\gamma}{d\theta} \\ y &= \sin(\theta)\gamma(\theta, T) + \cos(\theta)\frac{d\gamma}{d\theta} \end{aligned}$$

where $\gamma(\theta, T) = \gamma_0(\theta) - Ts(\theta, T)$ is the free energy per unit length of the interface. Deriving this is cool, but somewhat complicated.³⁰

(c) Wolff construction. Using the energy of eqn N.43 and approximating the entropy at low temperatures with the zero-temperature form eqn N.44, plot the Wulff shape for $T = 0.01, 0.1, 0.2, 0.4$, and $0.8J$ for one quadrant ($0 < \theta < \pi/2$). Hint: Ignore the parts of the face outside the quadrant; they are artifacts of the low temperature approximation. Does the shape become circular³¹ near $T_c = 2J/\log(1 + \sqrt{2}) \sim 2.269J$? Why not? If you’re ambitious, Rottman and Wortis’s article above gives the exact interfacial free energy for the 2D Ising model, which should fix this problem. What is the shape at low temperatures, where our approximation is good? Do we ever get true facets? The approximation of the interface as a staircase is a *solid-on-solid* model, which ignores overhangs. It is a good approximation at low temperatures.

²⁹If you put on an external field favoring up spins, then large clusters will grow and small clusters will shrink. The borderline cluster size is the *critical droplet* (see Entropy, Order Parameters, and Complexity, section 11.3). Indeed, the critical droplet will in general share the equilibrium crystal shape.

³⁰See Burton, Cabrera, and Frank, Appendix D. This is their equation D4, with $p \propto \gamma(\theta)$ given by equation D7.

³¹This is the emergent spherical symmetry at the Ising model critical point.

³²Flat regions demand that the free energy for adding a step onto the surface become

Our model does not describe faceting – one needs three dimensions to have a *roughening transition*, below which there are flat regions on the equilibrium crystal shape.³²

(N.21) **Condition Number and Accuracy.**³³ (Numerical) ③

You may think this exercise, with a 2x2 matrix, hardly demands a computer. However, it introduces tools for solving linear equations, condition numbers, singular value decomposition, all while illustrating subtle properties of matrix solutions. Use whatever linear algebra packages are provided in your software environment.

Consider the equation $\mathbf{Ax} = \mathbf{b}$, where

$$A = \begin{pmatrix} 0.780 & 0.563 \\ 0.913 & 0.659 \end{pmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} 0.217 \\ 0.254 \end{pmatrix}. \quad (\text{N.45})$$

The exact solution is $\mathbf{x} = (1, -1)$. Consider the two approximate solutions $\mathbf{x}_\alpha = (0.999, -1.001)$ and $\mathbf{x}_\beta = (0.341, -0.087)$.

(a) Compute the residuals \mathbf{r}_α and \mathbf{r}_β corresponding to the two approximate solutions. (The residual is $\mathbf{r} = \mathbf{b} - \mathbf{Ax}$.) Does the more accurate solution have the smaller residual?

(b) Compute the condition number³⁴ of A . Does it help you understand the result of (a)? (Hint: V^T maps the errors $\mathbf{x}_\alpha - \mathbf{x}$ and $\mathbf{x}_\beta - \mathbf{x}$ into what combinations of the two singular values?)

(c) Use a black-box linear solver to solve for \mathbf{x} . Subtract your answer from the exact one. Do you get within a few times the machine accuracy of 2.2×10^{-16} ? Is the problem the accuracy of the solution, or rounding errors in calculating A and \mathbf{b} ? (Hint: try calculating $\mathbf{Ax} - \mathbf{b}$.)

(N.22) **Sherman–Morrison formula.** (Numerical) ③

Consider the 5x5 matrices

$$T = \begin{pmatrix} E-t & t & 0 & 0 & 0 \\ t & E & t & 0 & 0 \\ 0 & t & E & t & 0 \\ 0 & 0 & t & E & t \\ 0 & 0 & 0 & t & E-t \end{pmatrix} \quad (\text{N.46})$$

and

$$C = \begin{pmatrix} E & t & 0 & 0 & t \\ t & E & t & 0 & 0 \\ 0 & t & E & t & 0 \\ 0 & 0 & t & E & t \\ t & 0 & 0 & t & E \end{pmatrix}. \quad (\text{N.47})$$

These matrices arise in one-dimensional models of crystals.³⁵ The matrix T is *tridiagonal*: its entries are zero except along the central diagonal and the entries neighboring the diagonal. Tridiagonal matrices are fast to solve; indeed, many routines will start by changing basis to make the array tridiagonal. The matrix C , on the other hand, has a nice periodic structure: each basis element has two neighbors, with the first and last basis elements now connected by t in the upper-right and lower-left corners. This periodic structure allows for analysis using Fourier methods (Bloch waves and \mathbf{k} -space).

For matrices like C and T which differ in only a few matrix elements³⁶ we can find C^{-1} from T^{-1} efficiently using the Sherman-Morrison formula (section 2.7).

Compute the inverse³⁷ of T for $E = 3$ and $t = 1$. Compute the inverse of C . Compare the differ-

infinite. Steps on the surfaces of three-dimensional crystals are long, and if they have positive energy per unit length the surface is faceted. To get an infinite energy for a step on a two-dimensional surface you need long-range interactions.

³³Adapted from Saul Teukolsky, 2003.

³⁴See section 2.6.2 for a technical definition of the condition number, and how it is related to singular value decomposition. Look on the Web for the more traditional definition(s), and how they are related to the accuracy.

³⁵As the Hamiltonian for electrons in a one-dimensional chain of atoms, t is the hopping matrix element and E is the on-site energy. As the potential energy for longitudinal vibrations in a one-dimensional chain, $E = -2t = K$ is the spring constant between two neighboring atoms. The tridiagonal matrix T corresponds to a kind of free boundary condition, while C corresponds in both cases to periodic boundary conditions.

³⁶More generally, this works whenever the two matrices differ by the outer product $\mathbf{u} \otimes \mathbf{v}$ of two vectors. By taking the two vectors to each have one non-zero component $u_i = u\delta_{ia}, v_j = v\delta_{jb}$, the matrices differ at one matrix element $\Delta_{ab} = uv$; for our matrices $\mathbf{u} = \mathbf{v} = 1, 0, 0, 0, 1$ (see section 2.7.2).

³⁷Your software environment should have a *solver* for tridiagonal systems, rapidly giving \mathbf{u} in the equation $T \cdot \mathbf{u} = \mathbf{r}$. It likely will not have a special routine for inverting tridiagonal matrices, but our matrix is so small it's not important.

ence $\Delta = T^{-1} - C^{-1}$ with that given by the Sherman-Morrison formula

$$\Delta = \frac{T^{-1} \mathbf{u} \otimes \mathbf{v} T^{-1}}{1 + \mathbf{v} \cdot T^{-1} \cdot \mathbf{u}}. \quad (\text{N.48})$$

(N.23) **Methods of interpolation.** (Numerical) ③

We've implemented four different interpolation methods for the function $\sin(x)$ on the interval $(-\pi, \pi)$. On the left, we see the methods using the five points $\pm\pi$, $\pm\pi/2$, and 0; on the right we see the methods using ten points. The graphs show the interpolation, its first derivative, its second derivative, and the error. The four interpolation methods we have implemented are (1) Linear, (2) Polynomial of degree three ($M = 4$), (3) Cubic spline, and (4) Barycentric rational interpolation. Which set of curves (A , B , C , or D) in Figure N.10 corresponds with which method?

(N.24) **Numerical definite integrals.** (Numerical) ③

In this exercise we will integrate the function you graphed in the first, warmup exercise:

$$F(x) = \exp(-6 \sin(x)). \quad (\text{N.49})$$

As discussed in Numerical Recipes, the word *integration* is used both for the operation that is inverse to differentiation, and more generally for finding solutions to differential equations. The old-fashioned term specific to what we are doing in this exercise is *quadrature*.

(a) *Black-box.* Using a professionally written black-box integration routine of your choice, integrate $F(x)$ between zero and π . Compare your answer to the analytic integral³⁸ ($\approx 0.34542493760937693$) by subtracting the analytic form from your numerical result. Read the documentation for your black box routine, and describe the combination of algorithms being used.

(b) *Trapezoidal rule.* Implement the trapezoidal rule with your own routine. Use it to calculate the same integral as in part (a). Calculate the estimated integral $\text{Trap}(h)$ for $N + 1$ points spaced at $h = \pi/N$, with $N = 1, 2, 4, \dots, 2^{10}$. Plot the estimated integral versus the spacing h . Does it extrapolate smoothly to the true value as $h \rightarrow 0$?

With what power of h does the error vanish? Replot the data as $\text{Trap}(h)$ versus h^2 . Does the error now vanish linearly?

Numerical Recipes tells us that the error is an even polynomial in h , so we can extrapolate the results of the trapezoidal rule using polynomial interpolation in powers of h^2 .

(c) *Simpson's rule (paper and pencil).* Consider a linear fit (i.e., $A + Bh^2$) to two points at $2h_0$ and h_0 on your $\text{Trap}(h)$ versus h^2 plot. Notice that the extrapolation to $h \rightarrow 0$ is A , and show that A is $4/3 \text{Trap}(h_0) - 1/3 \text{Trap}(2h_0)$. What is the net weight associated with the even points and odd points? Is this Simpson's rule?

(d) *Romberg integration.* Apply M -point polynomial interpolation (here extrapolation) to the data points $\{h^2, \text{Trap}(h)\}$ for $h = \pi/2, \dots, \pi/2^M$, with values of M between two and ten. (Note that the independent variable is h^2 .) Make a log-log plot of the absolute value of the error versus $N = 2^M$. Does this extrapolation improve convergence?

(e) *Gaussian Quadrature.* Implement Gaussian quadrature with N points optimally chosen on the interval $(0, \pi)$, with $N = 1, 2, \dots, 5$. (You may find the points and the weights appropriate for integrating functions on the interval $(-1, 1)$ on the course Web site; you will need to rescale them for use on $(0, \pi)$.) Make a log-log plot of the absolute value of your error as a function of the number of evaluation points N , along with the corresponding errors from the trapezoidal rule and Romberg integration.

(f) *Integrals of periodic functions.* Apply the trapezoidal rule to integrate $F(x)$ from zero to 2π , and plot the error on a log plot (log of the absolute value of the error versus N) as a function of the number of points N up to $N = 20$. (The true value should be around 422.44623805153909946.) Why does it converge so fast? (Hint: Don't get distracted by the funny alternation of accuracy between even and odd points.)

The location of the Gauss points depend upon the class of functions one is integrating. In part (e), we were using Gauss-Legendre quadrature, appropriate for functions which are analytic at the endpoints of the range of integration. In part (f), we have a function with *periodic boundary conditions*. For functions with

³⁸ $\pi(\text{BesselI}[0, 6] - \text{StruveL}[0, 6])$, according to Mathematica.

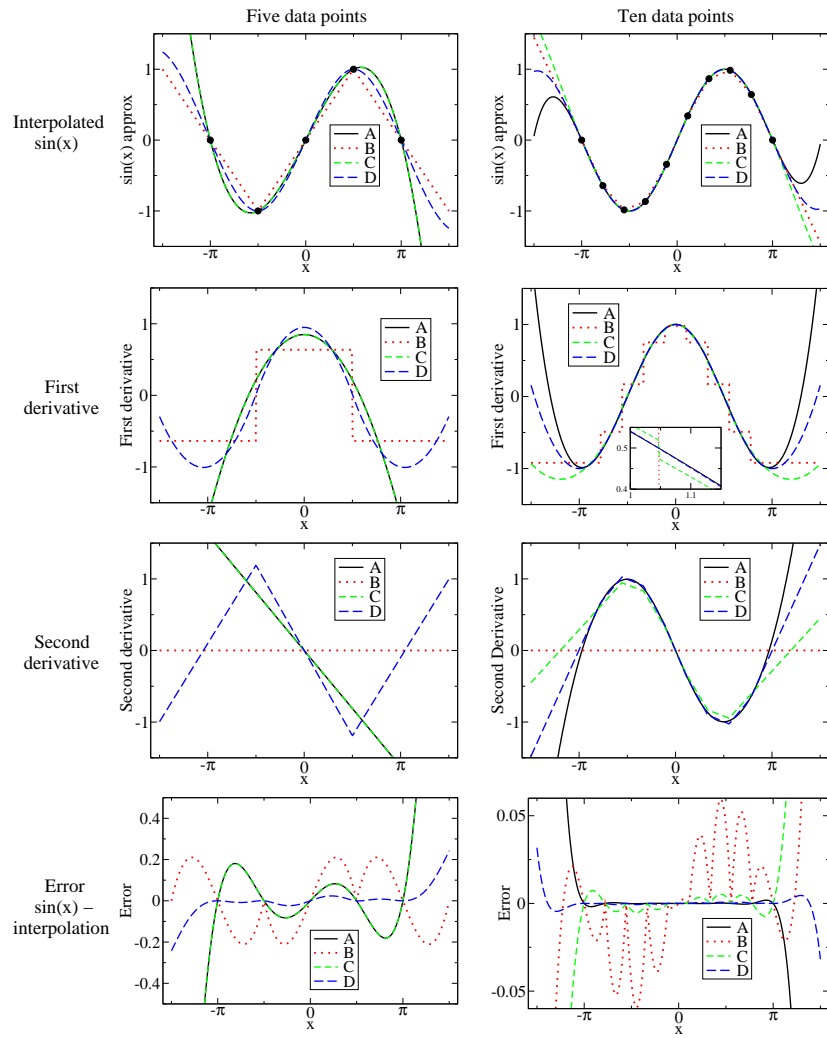


Fig. N.10 Interpolation methods.

periodic boundary conditions, the end-points are no longer special. What corresponds to Gaussian quadrature for periodic functions is just the trapezoidal rule: equally-weighted points at equally spaced intervals.

- (N.25) **Numerical derivatives.** (Rounding errors, Accuracy) ②

Calculate the numerical first derivative of the function $y(x) = \sin(x)$, using the centered two-point formula $dy/dx \sim (y(x+h) - y(x-h))/(2h)$, and plot the error $y'(x) - \cos(x)$ in the range $(-\pi, \pi)$ at 100 points. (Do a good job! Use the step-size h described in section 5.7 to optimize the sum of the truncation error and the rounding error. Also, make sure that the step-size h is exactly representable on the machine.) How does your actual error compare to the fractional error estimate given in section 5.7? Calculate and plot the numerical second derivative using the formula

$$\frac{d^2y}{dx^2} \sim \frac{y(x+h) - 2y(x) + y(x-h)}{h^2}, \quad (\text{N.50})$$

again optimizing h and making it exactly representable. Estimate your error again, and compare to the observed error.

- (N.26) **Summing series.** (Efficiency) ②
Write a routine to calculate the sum

$$s_n = \sum_{j=0}^n (-1)^j \frac{1}{2j+1}. \quad (\text{N.51})$$

As $n \rightarrow \infty$, $s_\infty = \pi/4$. About how many terms do you need to sum to get convergence to within 10^{-7} of this limit? Now try using Aitken's Δ^2 process to accelerate the convergence:

$$s'_n = s_n - \frac{(s_{n+1} - s_n)^2}{s_{n+2} - 2s_{n+1} + s_n}. \quad (\text{N.52})$$

About how many terms do you need with Aitken's method to get convergence to within 10^{-7} ?

- (N.27) **Random histograms.** (Random numbers) ②

(a) Investigate the random number generator for your system of choice. What is its basic algorithm? Its period?

(b) Plot a histogram with 100 bins, giving the normalized probability density of 100,000 random numbers sampled from (a) a uniform distribution in the range $0 < x < 2\pi$, (b) an exponential distribution $\rho(x) = 6 \exp(-6x)$, and (c) a normal distribution of mean $\bar{x} = 3\pi/2$ and

standard deviation $\sigma = 1/\sqrt{6}$. Before each plot, set the seed of your random number generator. Do you now get the same plot when you repeat?

- (N.28) **Monte Carlo integration.** (Random numbers, Robust algorithms) ③

How hard can numerical integration be? Suppose the function f is wildly nonanalytic, or has a peculiar or high-dimensional domain of integration? In the worst case, one can always try Monte Carlo integration. The basic idea is to pepper points at random in the integration interval. The integration volume times the average of the function $V\langle f \rangle$ is the estimate of the integral.

As one might expect, the expected error in the integral after N evaluations is given by $1/\sqrt{N-1}$ times the standard deviation of the sampled points (NR equation 7.7.1).

(a) Monte Carlo in one dimensional integrals. Use Monte Carlo integration to estimate the integral of the function introduced in the preliminary exercises

$$y(x) = \exp(-6 \sin(x)) \quad (\text{N.53})$$

over $0 \leq x < 2\pi$. (The correct value of the integral is around 422.446.) How many points do you need to get 1% accuracy? Answer this last question both by experimenting with different random number seeds, and by calculating the expected number from the standard deviation. (You may use the fact that $\langle y^2 \rangle = (1/2\pi) \int_0^{2\pi} y^2(x) dx \approx 18948.9$.)

Monte Carlo integration is not the most efficient method for calculating integrals of smooth functions like $y(x)$. Indeed, since $y(x)$ is periodic in the integration interval, equally spaced points weighted equally (the trapezoidal rule) gives exponentially rapid convergence; it takes only nine points to get 1% accuracy. Even for smooth functions, though, Monte Carlo integration is useful in high dimensions.

(b) Monte Carlo in many dimensions (No detailed calculations expected). For a hypothetical ten-dimensional integral, if we used a regular grid with nine points along each axis, how many function evaluations would we need for equivalent accuracy? Does the number of Monte Carlo points needed depend on the dimension of the space, presuming (perhaps naively) that the variance of the function stays fixed?

Our function $y(x)$ is quite close to a Gaussian. (Why? Taylor expand $\sin(x)$ about $x = 3\pi/2$.)

We can use this to do *importance sampling*. The idea is to evaluate the integral of $h(x)g(x)$ by randomly sampling h with probability g , picking $h(x) = y(x)/g(x)$. The variance is then $\langle h^2 \rangle - \langle h \rangle^2$. In order to properly sample the tail near $x = \pi/2$, we should mix a Gaussian and a uniform distribution:

$$g(x) = \frac{\epsilon}{2\pi} + \frac{1-\epsilon}{\sqrt{\pi/3}} \exp(-6(x - 3\pi/2)^2/2). \quad (\text{N.54})$$

I found minimum variance around $\epsilon = 0.005$.

(c) Importance Sampling (Optional for 480). Generate 1000 random numbers with probability distribution g .³⁹ Use these to estimate the integral of $y(x)$. How accurate is your answer?

(N.29) **The Birthday Problem.** (Sorting, Random numbers) ②

Remember birthday parties in your elementary school? Remember those years when two kids had the same birthday? How unlikely!

How many kids would you need in class to get, more than half of the time, at least two with the same birthday?

(a) Numerical. Write a routine `BirthdayCoincidences(K, C)` that returns the fraction among C classes for which at least two kids (among K kids per class) have the same birthday. (Hint: By sorting a random list of integers, common birthdays will be adjacent.) Plot this probability versus K for a reasonably large value of C . Is it a surprise that your classes had overlapping birthdays when you were young?

One can intuitively understand this, by remembering that to avoid a coincidence there are $K(K-1)/2$ pairs of kids, all of whom must have different birthdays (probability $364/365 = 1 - 1/D$, with D days per year).

$$P(K, D) \approx (1 - 1/D)^{K(K-1)/2} \quad (\text{N.55})$$

This is clearly a crude approximation – it doesn't vanish if $K > D$! Ignoring subtle correlations, though, it gives us a net probability

$$\begin{aligned} P(K, D) &\approx \exp(-1/D)^{K(K-1)/2} \\ &\approx \exp(-K^2/(2D)) \end{aligned} \quad (\text{N.56})$$

Here we've used the fact that $1 - \epsilon \approx \exp(-\epsilon)$, and assumed that K/D is small.

(b) Analytical. Write the exact formula giving the probability, for K random integers among D choices, that no two kids have the same birthday. (Hint: What is the probability that the second kid has a different birthday from the first? The third kid has a different birthday from the first two?) Show that your formula does give zero if $K > D$. Converting the terms in your product to exponentials as we did above, show that your answer is consistent with the simple formula above, if $K \ll D$. Inverting equation N.56, give a formula for the number of kids needed to have a 50% chance of a shared birthday.

Some years ago, we were doing a large simulation, involving sorting a lattice of 1000^3 random fields (roughly, to figure out which site on the lattice would trigger first). If we want to make sure that our code is unbiased, we want different random fields on each lattice site – a giant birthday problem.

Old-style random number generators generated a random integer (2^{32} 'days in the year') and then divided by the maximum possible integer to get a random number between zero and one. Modern random number generators generate all 2^{52} possible doubles between zero and one.

(c) If there are $2^{31} - 1$ distinct four-byte positive integers, how many random numbers would one have to generate before one would expect coincidences half the time? Generate lists of that length, and check your assertion. (Hints: It is faster to use array operations, especially in interpreted languages. I generated a random array with N entries, sorted it, subtracted the first $N-1$ entries from the last $N-1$, and then called `min` on the array.) Will we have to worry about coincidences with an old-style random number generator? How large a lattice $L \times L \times L$ of random double precision numbers can one generate with modern generators before having a 50% chance of a coincidence? If you have a fast machine with a large memory, you might test this too.

³⁹Take $(1 - \epsilon)M$ Gaussian random numbers and ϵM random numbers uniformly distributed on $(0, 2\pi)$. The Gaussian has a small tail that extends beyond the integration range $(0, 2\pi)$, so the normalization of the second term in the definition of g is not quite right. You can fix this by simply throwing away any samples that include points outside the range.

(N.30) **Dashboard Potential.** (Solving) ②

Consider a dashboard potential⁴⁰

$$V(r) = A_1 \cos(r) + A_2 \cos(2r) - Fr \quad (\text{N.57})$$

with $A_1 = 5$, $A_2 = 1$, and F initially equal to 1.5.

(a) Plot $V(r)$ over $(-10, 10)$. Numerically find the local maximum of V near zero, and the local minimum of V to the left (negative side) of zero. What is the potential energy barrier for moving from one well to the next in this potential?

Usually finding the minimum is only a first step – one wants to explore how the minimum moves and disappears. . .

(b) Increasing the external tilting field F , graphically roughly locate the field F_c where the barrier disappears, and the location r_c at this field where the potential minimum and maximum merge. (This is a saddle-node bifurcation.) Give the criterion on the first derivative and the second derivative of $V(r)$ at F_c and r_c . Using these two equations, numerically use a root-finding routine to locate the saddle-node bifurcation F_c and r_c .

(N.31) **Sloppy Minimization.** (Minimization) ③

“With four parameters I can fit an elephant. With five I can make it waggle its trunk.” This statement, attributed to many different sources (from Carl Friedrich Gauss to Fermi), reflects the problems found in fitting multiparameter models to data. One almost universal problem is *sloppiness* – the parameters in the model are poorly constrained by the data.⁴¹

Consider the classic ill-conditioned problem of fitting exponentials to radioactive decay data. If you know that at $t = 0$ there are equal quantities of N radioactive materials with half-lives γ_n , the radioactivity that you would measure is

$$y(t) = \sum_{n=0}^{N-1} \gamma_n \exp(-\gamma_n t). \quad (\text{N.58})$$

Now, suppose you don’t know the decay rates γ_n . Can you reconstruct them by fitting the data to experimental data $y_0(t)$?

Let’s consider the problem with two radioactive decay elements $N = 2$. Suppose the actual decay constants for $y_0(t)$ are $\alpha_n = n + 2$ (so the experiment has $\gamma_0 = \alpha_0 = 2$ and $\gamma_1 = \alpha_1 = 3$), and we try to minimize the least-squared error⁴² in the fits C :

$$C[\gamma] = \int_0^\infty (y(t) - y_0(t))^2 dt. \quad (\text{N.59})$$

You can convince yourself that the least-squared error is

$$\begin{aligned} C[\gamma] &= \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} \left(\frac{\gamma_n \gamma_m}{\gamma_n + \gamma_m} \right. \\ &\quad \left. + \frac{\alpha_n \alpha_m}{\alpha_n + \alpha_m} - 2 \frac{\gamma_n \alpha_m}{\gamma_n + \alpha_m} \right) \\ &= \frac{49}{10} + \frac{\gamma_0}{2} - \frac{4\gamma_0}{2 + \gamma_0} - \frac{6\gamma_0}{3 + \gamma_0} \\ &\quad + \frac{\gamma_1}{2} - \frac{4\gamma_1}{2 + \gamma_1} - \frac{6\gamma_1}{3 + \gamma_1} + \frac{2\gamma_0 \gamma_1}{\gamma_0 + \gamma_1}. \end{aligned}$$

(a) Draw a contour plot of C in the square $1.5 < \gamma_n < 3.5$, with enough contours (perhaps non-equally spaced) so that the two minima can be distinguished. (You’ll also need a fairly fine grid of points.)

One can see from the contour plot that measuring the two rate constants separately would be a challenge. This is because the two exponentials have similar shapes, so increasing one decay rate and decreasing the other can almost perfectly compensate for one another.

(b) If we assume both elements decay with the same decay constant $\gamma = \gamma_0 = \gamma_1$, minimize the cost to find the optimum choice for γ . Where is this point on the contour plot? Plot $y_0(t)$ and $y(t)$ with this single-exponent best fit on the same graph, over $0 < t < 2$. Do you agree that it would be difficult to distinguish these two fits?

This problem can become much more severe in higher dimensions. The banana-shaped ellipses in your contour plot can become needle-like, with aspect ratios of more than a thousand to one (about the same as a human hair). Following these thin paths down to the true minima can

⁴⁰A washboard is what people used to hand-wash clothing. It is held at an angle, and has a series of corrugated ridges; one holds the board at an angle and rubs the wet clothing on it. Washboard potentials arise in the theory of superconducting Josephson junctions, in the motion of defects in crystals, and in many other contexts.

⁴¹‘Elephantness’ says that a wide range of behaviors can be exhibited by varying the parameters over small ranges. ‘Sloppiness’ says that a wide range of parameters can yield approximately the same behavior. Both reflect a *skewness* in the relation between parameters and model behavior.

⁴²We’re assuming perfect data at all times, with uniform error bars.

be a challenge for multidimensional minimization programs.

(c) Find a method for storing and plotting the locations visited by your minimization routine (values of (γ_0, γ_1) at which it evaluates C while searching for the minimum). Starting from $\gamma_0 = 1, \gamma_1 = 4$, minimize the cost using as many methods as is convenient within your programming environment, such as Nelder-Mead, Powell, Newton, Quasi-Newton, conjugate gradient, Levenberg-Marquardt (nonlinear least squares)... (Try to use at least one that does not demand derivatives of the function). Plot the evaluation points on top of the contour plot of the cost for each method. Compare the number of function evaluations needed for the different methods.

(N.32) **Sloppy Monomials.**⁴³ (Eigenvalues, Fitting) ③

We have seen that the same function $f(x)$ can be approximated in many ways. Indeed, the same function can be fit in the same interval by the same type of function in several different ways! For example, in the interval $[0, 1]$, the function $\sin(2\pi x)$ can be approximated (badly) by a fifth-order Taylor expansion, a Chebyshev polynomial, or a least-squares (Legendre⁴⁴) fit:

$$\begin{aligned} f(x) &= \sin(2\pi x) \\ &\approx 0.000 + 6.283x + 0.000x^2 - 41.342x^3 \\ &\quad + 0.000x^4 + 81.605x^5 \quad \text{Taylor} \\ &\approx 0.007 + 5.652x + 9.701x^2 - 95.455x^3 \\ &\quad + 133.48x^4 - 53.39x^5 \quad \text{Chebyshev} \\ &\approx 0.016 + 5.410x + 11.304x^2 - 99.637x^3 \\ &\quad + 138.15x^4 - 55.26x^5 \quad \text{Legendre} \end{aligned}$$

It is not a surprise that the best fit polynomial differs from the Taylor expansion, since the latter is not a good approximation. But it is a surprise that the last two polynomials are so different. The maximum error for Legendre is less than 0.02, and for Chebyshev is less than 0.01,

even though the two polynomials differ by

$$\begin{aligned} \text{Chebyshev} - \text{Legendre} &= \quad \quad \quad \text{(N.60)} \\ &- 0.009 + 0.242x - 1.603x^2 \\ &+ 4.182x^3 - 4.67x^4 + 1.87x^5 \end{aligned}$$

a polynomial with coefficients two hundred times larger than the maximum difference!

This flexibility in the coefficients of the polynomial expansion is remarkable. We can study it by considering the dependence of the quality of the fit on the parameters. Least-squares (Legendre) fits minimize a cost C^{Leg} , the integral of the squared difference between the polynomial and the function:

$$C^{\text{Leg}} = (1/2) \int_0^1 (f(x) - \sum_{m=0}^M a_m x^m)^2 dx. \quad \text{(N.61)}$$

How quickly does this cost increase as we move the parameters a_m away from their best-fit values? Varying any one monomial coefficient will of course make the fit bad. But apparently certain coordinated changes of coefficients don't cost much – for example, the difference between least-squares and Chebyshev fits given in eqn (N.60).

How should we explore the dependence in arbitrary directions in parameter space? We can use the eigenvalues of the Hessian to see how sensitive the fit is to moves along the various eigenvectors...

(a) Note that the first derivative of the cost C^{Leg} is zero at the best fit. Show that the Hessian second derivative of the cost is

$$H_{mn}^{\text{Leg}} = \frac{\partial^2 C^{\text{Leg}}}{\partial a_m \partial a_n} = \frac{1}{m+n+1}. \quad \text{(N.62)}$$

This Hessian is the Hilbert matrix, famous for being ill-conditioned (having a huge range of eigenvalues). Tiny eigenvalues of H^{Leg} correspond to directions in polynomial space where the fit doesn't change.

(b) Calculate the eigenvalues of the 6×6 Hessian for fifth-degree polynomial fits. Do they span a large range? How big is the condition number?

(c) Calculate the eigenvalues of larger Hilbert matrices. At what size do your eigenvalues seem

⁴³Thanks to Joshua Waterfall, whose research is described here.

⁴⁴The orthogonal polynomials used for least-squares fits on $[-1,1]$ are the Legendre polynomials, assuming continuous data points. Were we using orthogonal polynomials for this exercise, we would need to shift them for use in $[0,1]$.

contaminated by rounding errors? Plot them, in order of decreasing size, on a semi-log plot to illustrate these rounding errors.

Notice from Eqn N.62 that the dependence of the polynomial fit on the monomial coefficients is *independent of the function $f(x)$* being fitted. We can thus vividly illustrate the sloppiness of polynomial fits by considering fits to the *zero function* $f(x) \equiv 0$. A polynomial given by an eigenvector of the Hilbert matrix with small eigenvalue must stay close to zero everywhere in the range $[0, 1]$. Let's check this.

(d) Calculate the eigenvector corresponding to the smallest eigenvalue of H^{Leg} , checking to make sure its norm is one (so the coefficients are of order one). Plot the corresponding polynomial in the range $[0, 1]$: does it stay small everywhere in the interval? Plot it in a larger range $[-1, 2]$ to contrast its behavior inside and outside the fit interval.

This turns out to be a fundamental property that is shared with many other multiparameter fitting problems. Many different terms are used to describe this property. The fits are called *ill-conditioned*: the parameters a_n are not well constrained by the data. The *inverse problem* is challenging: one cannot practically extract the parameters from the behavior of the model. Or, as our group describes it, the fit is *sloppy*: only a few directions in parameter space (eigenvectors corresponding to the largest eigenvalues) are constrained by the data, and there is a huge space of models (polynomials) varying along sloppy directions that all serve well in describing the data.

At root, the problem with polynomial fits is that all monomials x^n have similar shapes on $[0, 1]$: they all start flat near zero and bend upward. Thus they can be traded for one another; the coefficient of x^4 can be lowered without changing the fit if the coefficients of x^3 and x^5 are suitably adjusted to compensate. Indeed, if we change basis from the coefficients a_n of the monomials x^n to the coefficients ℓ_n of the orthogonal (shifted Legendre) polynomials, the situation completely changes. The Legendre polynomials are designed to be different in shape (orthogonal), and hence cannot be traded for one another. Their coefficients ℓ_n are thus well de-

termined by the data, and indeed the Hessian for the cost C^{Leg} in terms of this new basis is the identity matrix.

Numerical Recipes states a couple of times that using equally-spaced points gives ill-conditioned fits. Will the Chebyshev fits, which emphasize the end-points of the interval, give less sloppy coefficients? The Chebyshev polynomial for a function f (in the limit where many terms are kept) minimizes a different cost: the squared difference weighted by the extra factor $1/\sqrt{x(1-x)}$:

$$C^{\text{Cheb}} = \int_0^1 \frac{(f(x) - \sum_{m=0}^M c_m x^m)^2}{\sqrt{x(1-x)}} dx. \quad (\text{N.63})$$

One can show that the Hessian giving the dependence of C^{Cheb} on c_m is

$$H_{mn}^{\text{Cheb}} = \frac{\partial^2 C^{\text{Cheb}}}{\partial c_m \partial c_n} = \frac{2^{1-2(m+n)} \pi (2(m+n) - 1)!}{(m+n-1)!(m+n)!}. \quad (\text{N.64})$$

with $H_{00}^{\text{Cheb}} = \pi$ (doing the integral explicitly, or taking the limit $m, n \rightarrow 0$).

(d) (Optional for 480) Calculate the eigenvalues of H^{Cheb} for fifth-degree polynomial fits.

So, sloppiness is not peculiar to least-squares fits; Chebyshev polynomial coefficients are sloppy too.

(N.33) **Conservative differential equations: Accuracy and fidelity.** (Ordinary differential equations) ③

In this exercise, we will solve for the motion of a particle of mass $m = 1$ in the potential

$$V(y) = (1/8) y^2 (-4A^2 + \log^2(y^2)). \quad (\text{N.65})$$

That is,

$$\begin{aligned} d^2 y / dt^2 &= -dV/dy \\ &= -(1/4)y(-4A^2 + 2 \log(y^2) + \log^2(y^2)). \end{aligned} \quad (\text{N.66})$$

We will start the particle at $y_0 = 1$, $v_0 = dy/dt|_0 = -A$, and choose $A = 6$.

(a) Show that the solution to this differential equation is⁴⁵

$$F(t) = \exp(-6 \sin(t)). \quad (\text{N.67})$$

Note that the potential energy $V(y)$ is zero at the five points $y = 0$, $y = \pm \exp(\pm A)$.

(b) Plot the potential energy for $-3/2 \exp(\pm A) < y < 3/2 \exp(\pm A)$ (both zoomed

⁴⁵I worked backward to do this. I set the kinetic energy to $1/2(dF/dt)^2$, set the potential energy to minus the kinetic energy, and then substituted y for t by solving $y = F(t)$.

in near $y = 0$ and zoomed out). The correct trajectory should oscillate in the potential well with $y > 0$, turning at two points whose energy is equal to the initial total energy. *What is this initial total energy for our the particle? How much of an error in the energy would be needed, for the particle to pass through the origin when it returns? Compare this error to the maximum kinetic energy of the particle (as it passes the bottom of the well).* Small energy errors in our integration routine can thus cause significant changes in the trajectory.

(c) (Black-box) *Using a professionally written black-box differential equation solver of your choice, solve for $y(t)$ between zero and 4π , at high precision. Plot your answer along with $F(t)$ from part (a). (About half of the solvers will get the answer qualitatively wrong, as expected from part (b).) Expand your plot to examine the region $0 < y < 1$; is energy conserved? Finally, read the documentation for your black box routine, and describe the combination of algorithms being used. You may wish to implement and compare more than one algorithm, if your black-box has that option.*

Choosing an error tolerance for your differential equation limits the error in each time step. If small errors in early time steps lead to important changes later, your solution may be quite different from the correct one. Chaotic motion, for example, can never be accurately simulated on a computer. All one can hope for is a *faithful* simulation – one where the motion is qualitatively similar to the real solution. Here we find an important discrepancy – the energy of the numerical solution is drifting upward or downward, where energy should be exactly conserved in the true solution. Here we get a dramatic change in the trajectory for a small quantitative error, but any drift in the energy is qualitatively incorrect. The Leapfrog algorithm is a primitive looking method for solving for the motion of a particle in a potential. It calculates the next position from the previous two:

$$y(t+h) = 2y(t) - y(t-h) + h^2 f[y(t)] \quad (\text{N.68})$$

where $f[y] = -dV/dy$. Given an initial position $y(0)$ and an initial velocity $v(0)$, one can initial-

ize and finalize

$$\begin{aligned} y(h) &= y(0) + h(v(0) + (h/2)f[y(0)]) \\ v(T) &= (y(T) - y(T-h))/h + (h/2)f[y(T)] \end{aligned} \quad (\text{N.69})$$

Leapfrog is one of a variety of *Verlet* algorithms. In a more general context, this is called *Störmer's rule*, and can be extrapolated to zero step-size h as in the Bulirsch-Stoer algorithms. A completely equivalent algorithm, with more storage but less roundoff error, is given by computing the velocity v at the midpoints of each time step:

$$\begin{aligned} v(h/2) &= v(0) + (h/2)f[y(0)] \\ y(h) &= y(0) + h v(h/2) \\ v(t+h/2) &= v(t-h/2) + hf[y(t)] \\ y(t+h) &= y(t) + h v(t+h/2) \end{aligned} \quad (\text{N.70})$$

where we may reconstruct $v(t)$ at integer time steps with $v(t) = v(t-h/2) + (h/2)f[y(t)]$.

(d) (Leapfrog and symplectic methods) *Show that $y(t+h)$ and $y(h)$ from eqns N.68 and N.69 converge to the true solutions as $h \rightarrow 0$ – that is, the time-step error compared to the solution of $d^2y/dt^2 = f[y]$ vanishes faster than h . To what order in h are they accurate after one time step?⁴⁶ Implement Leapfrog, and apply it to solving equation N.66 in the range $0 < x < 4\pi$, starting with step size $h = 0.01$. How does the accuracy compare to your more sophisticated integrator, for times less than $t = 2$? Zoom in to the range $0 < y < 1$, and compare with your packaged integrator and the true solution. Which has the more accurate period (say, by measuring the time to the first minimum in y)? Which has the smaller energy drift (say, by measuring the change in depth between subsequent minima)?*

Fidelity is often far more important than accuracy in numerical simulations. Having an algorithm that has a small time-step error but gets the behavior qualitatively wrong is less useful than a cruder answer that is faithful to the physics. Leapfrog here is capturing the oscillating behavior far better than vastly more sophisticated algorithms, even though it gets the period wrong.

How does Leapfrog do so well? Systems like particle motion in a potential are Hamiltonian

⁴⁶Warning: for subtle reasons, the errors in Leapfrog apparently build up quadratically as one increases the number of time steps, so if your error estimate after one time step is h^n the error after $N = T/h$ time steps can't be assumed to be $Nh^n \sim h^{n-1}$, but is actually $N^2h^n \sim h^{N-2}$.

systems. They not only conserve energy, but they also have many other striking properties like conserving phase-space volume (Liouville's theorem, the basis for statistical mechanics). Leapfrog, in disguise, is also a Hamiltonian system. (Eqns N.70 can be viewed as a composition of two canonical transformations – one advancing the velocities at fixed positions, and one advancing the positions at fixed velocities.) Hence it exactly conserves an approximation to the energy – and thus doesn't suffer from energy drift, satisfies Liouville's theorem, etc. Leapfrog and

the related Verlet algorithm are called *symplectic* because they conserve the *symplectic form* that mathematicians use to characterize Hamiltonian systems.

It is often vastly preferable to do an exact simulation (apart from rounding errors) of an approximate system, rather than an approximate analysis of an exact system. That way, one can know that the results will be physically sensible (and, if they are not, that the bug is in your model or implementation, and not a feature of the approximation).

References

- [1] Buchel, A. and Sethna, J. P. (1996). Elastic theory has zero radius of convergence. *Physical Review Letters*, **77**, 1520.
- [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.
- [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.
- [4] Dyson, Freeman (1960). Search for artificial stellar sources of infrared radiation. *Science*, **131**, 1667–1668.
- [5] Mézard, M. and Montanari, A. (To be published (2006)). *Constraint satisfaction networks in Physics and Computation*. Clarendon Press, Oxford.
- [6] Otani, N. (2004). Home page. <http://otani.vet.cornell.edu>.
- [7] Sethna, J. P. (1997). Quantum electrodynamics has zero radius of convergence. <http://www.lasp.cornell.edu/sethna/Cracks/QED.html>.
- [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>.
- [9] Skriver, H. (2004). The hls metals database. <http://databases.fysik.dtu.dk/hlsPT/>.
- [10] Smith, R. L. (To be published (2006)). *Environmental Statistics*.
- [11] Winfree, A. T. (1991). Varieties of spiral wave behavior: An experimentalist’s approach to the theory of excitable media. *Chaos*, **1**, 303–34.
- [12] Zinn-Justin, J. (1996). *Quantum field theory and critical phenomena (3rd edition)*. Oxford University Press.