

Entropy, Order Parameters, and Complexity: Solutions to Exercises

Stephen Hicks, Bruno Rousseau, Nick Taylor, and James P. Sethna

Contents

What is Statistical Mechanics	4
1.1 Quantum Dice	4
1.3 Waiting Times	4
1.6 Random Matrix Theory	6
Random walks and emergent properties	11
2.1 Random Walks in Grade Space	11
2.2 Photon diffusion in the Sun	11
2.5 Generating Random Walks	11
2.6 Fourier and Green	12
2.8 Polymers and Random Walks	12
2.11 Stocks, Volatility and Diversification	17
2.12 Computational Finance: Pricing Derivatives	19
Temperature and equilibrium	23
3.5 Hard Sphere Gas	23
3.6 Connecting Two Macroscopic Systems	24
3.8 Microcanonical Energy Fluctuations	25
3.9 Gauss and Poisson	25
3.10 Triple Product Relation	28
3.11 Maxwell Relations	29
Phase-space dynamics and ergodicity	29
4.2 Liouville vs. the damped pendulum	29
4.3 Invariant Measures	30
4.4 Jupiter! and the KAM Theorem	32
Entropy	36
5.1 Life and the Heat Death of the Universe	36
5.2 Burning Information and Maxwellian Demons	37
5.3 Reversible Computation	37
5.4 Black Hole Thermodynamics	38
5.5 P-V Diagram	39
5.6 Carnot Refrigerator	40
5.7 Does Entropy Increase?	40
5.8 The Arnol'd Cat	41
5.9 Chaos, Lyapunov, and Entropy Increase	42
5.10 Entropy Increases: Diffusion	42
5.11 Entropy of Glasses	42
5.12 Rubber Band	44
5.13 How Many Shuffles?	46
5.15 Shannon entropy	47
5.17 Deriving Entropy	48
Free energy	50

6.3 Negative Temperature	50
6.4 Molecular Motors: Which Free Energy?	53
6.5 Laplace	55
6.7 Legendre	55
6.8 Euler	56
6.9 Gibbs-Duhem	57
6.10 Clausius-Clapeyron	57
6.11 Barrier Crossing	58
6.13 Pollen and Hard Squares	59
6.14 Statistical Mechanics and Statistics	61
Quantum statistical mechanics	61
7.1 Ensembles and quantum statistics	61
7.2 Phonons and Photons are Bosons	65
7.3 Phase Space Units and the Zero of Entropy	66
7.4 Does Entropy Increase in Quantum Systems?	67
7.5 Photon Density Matrices	68
7.6 Spin Density Matrix	69
7.8 Einstein's A and B	69
7.9 Bosons are Gregarious: Superfluids and Lasers	71
7.10 Crystal Defects	73
7.11 Phonons on a String	74
7.12 Semiconductors	75
7.13 Bose Condensation in a Band	76
7.15 The Photon-dominated Universe	77
7.16 White Dwarves, Neutron Stars, and Black Holes	79
Calculation and computation	80
8.2 Ising Fluctuations and Susceptibilities	80
8.3 Waiting for Godot, and Markov	82
8.4 Red and Green Bacteria	83
8.5 Detailed Balance	85
8.6 Metropolis	86
8.8 Wolff	86
8.10 Stochastic Cells	87
8.12 Entropy Increases! Markov Chains	88
Order parameters, broken symmetry, and topology	89
9.1 Topological Defects in Nematic Liquid Crystals	89
9.2 Topological Defects in the XY Model	90
9.3 Defect energetics and Total divergence terms	91
9.4 Domain Walls in Magnets	91
9.5 Landau Theory for the Ising Model	93
9.6 Symmetries and Wave Equations	94
9.7 Superfluid Order and Vortices	95
9.8 Superfluids: Density Matrices and ODLRO	96
Correlations, response, and dissipation	99
10.1 Microwave Background Radiation	99
10.2 Pair distributions and molecular dynamics	100
10.3 Damped Oscillators	102
10.4 Spin	106
10.5 Telegraph Noise in Nanojunctions	108
10.6 Fluctuations-Dissipation: Ising	109

10.7 Noise and Langevin equations	110
10.8 Magnet Dynamics	111
10.9 Quasiparticle poles and Goldstone's theorem	115
Abrupt phase transitions	116
11.1 Maxwell and Van Der Waals	116
11.4 Nucleation in the Ising Model	117
11.5 Nucleation of Dislocation Pairs	118
11.5 Nucleation of Dislocation Pairs	119
11.6 Coarsening in the Ising Model	120
11.7 Origami Microstructures	120
11.8 Minimizing Sequences	122
11.9 Snowflakes and Linear Stability	123
Continuous phase transitions	124
12.2 Scaling and corrections to scaling	124
12.3 Scaling and Coarsening	124
12.4 Bifurcation Theory and Phase Transitions	124
12.5 Mean-field theory	125
12.7 Renormalization Group Trajectories	126
12.8 Superconductivity and the Renormalization Group	127
12.10 Renormalization Group and the Central Limit Theorem (Short)	128
12.11 Renormalization Group and the Central Limit Theorem (Long)	128
12.13 Hysteresis Model: Scaling and Exponent Equalities	131

1.1 Quantum Dice.

(a) Presume the dice are fair: each of the three numbers of dots shows up $1/3$ of the time. For a legal turn rolling a die twice on *Bosons*, what is the probability $\rho(4)$ of rolling a 4? Similarly, among the legal *Fermion* turns rolling two dice, what is the probability $\rho(4)$?

The probability of rolling a four in *Bosons* or *Fermions* is given by

$$\text{probability} = \frac{\text{number of legal rolls giving four}}{\text{total number of legal rolls}}. \quad (1)$$

From figure 1.4 in the homework we can count off the appropriate number of rolls to find:

$$\rho_{\text{Bosons}}(4) = \rho_{\text{Fermions}}(4) = 1/3.$$

(b) For a legal turn rolling three 'three-sided' dice in *Fermions*, what is the probability $\rho(6)$ of rolling a 6?

For a legal roll in *Fermions* the dice are not allowed to show a particular number more than once, so in rolling three dice there is only one possible legal roll—1, 2, 3. The probability of rolling a 6 is therefore one: $\rho_{\text{Fermions}}(6) = 1$.

(c) In a turn of three rolls, what is the enhancement of probability of getting triples in *Bosons* over that in *Distinguishable*? In a turn of M rolls, what is the enhancement of probability for generating an M -tuple (all rolls having the same number of dots showing)?

There are exactly three legal rolls that are triples in either *Bosons* or *Distinguishable*—(1,1,1), (2,2,2) and (3,3,3). The total number of legal rolls in *Bosons* rolling three dice is $\binom{5}{3} = 10$, while in *Distinguishable* it is $3^3 = 27$. Thus, the enhancement of probability of getting triples in three rolls in *Bosons* over that in *Distinguishable* is

$$\frac{\rho_{\text{Bosons}}(\text{triples})}{\rho_{\text{Dist.}}(\text{triples})} = \frac{3/10}{3/27} = \frac{27}{10}.$$

For the general case of M rolls generating an M -tuple with three-sided dice, the enhancement of probability is

$$\frac{\rho_{\text{Bosons}}(M\text{-tuple})}{\rho_{\text{Dist.}}(M\text{-tuple})} = \frac{3/\binom{M+2}{M}}{3/3^M} = \frac{2 \cdot 3^M}{(M+2)(M+1)}.$$

and we can check that this agrees with the above for $M = 3$. The general solution for N -sided dice is

$$\frac{\rho_{\text{Bosons}}(M\text{-tuple}, N)}{\rho_{\text{Dist.}}(M\text{-tuple}, N)} = \frac{N^{M-1} M! N!}{(N+M-1)!}.$$

1.3 Waiting Times.

(a) Verify that each hour the average number of cars passing the observer is 12.

We have $\tau = 5\text{min}$ and a probability $\frac{dt}{\tau}$ of a car passing in the time dt . We integrate

$$\langle N \rangle = \int_0^T \frac{dt}{\tau} = \frac{T}{\tau} = 12$$

for $T = 60\text{min}$.

(b) What is the probability $P_{\text{bus}}(n)$ that n buses pass the observer in a randomly chosen 10min interval? And what is the probability $P_{\text{car}}(n)$ that n cars pass the observer in the same time interval?

Since buses come regularly every 5min, the number of buses passing in an interval depends only on when the interval starts. Unless the interval starts exactly as a bus passes, the observer will count two buses. But there is an infinitesimal chance of the interval starting exactly then, so that

$$P_{\text{bus}}(n) = \begin{cases} 1 & n = 2 \\ 0 & \text{otherwise} \end{cases}.$$

For cars, we break the $T = 10\text{min}$ interval into $N = \frac{T}{dt}$ chunks of length dt . In any given chunk, the probability of car passing is $\frac{dt}{\tau}$ and thus the probability of no car passing is $1 - \frac{dt}{\tau}$. For n cars to pass, we need exactly n chunks with cars and $N - n \approx N$ chunks without cars ($N \gg n$). The n chunks with cars can be arranged in any of $\binom{N}{n} \approx \frac{N^n}{n!}$ orderings, so that

$$P_{\text{car}}(n) = \lim_{dt \rightarrow 0} \frac{N^n}{n!} \left(1 - \frac{dt}{\tau}\right)^{T/dt} \left(\frac{dt}{\tau}\right)^n = \frac{1}{n!} \left(\frac{T}{\tau}\right)^n e^{-T/\tau}.$$

(c) What is the probability distribution ρ_{bus} and ρ_{car} for the time interval Δ between two successive buses and cars, respectively? What are the means of these distributions?

The interval between buses is always τ , so that the distribution is given by a Dirac delta function:

$$\rho_{\text{bus}}^{\text{gap}}(\Delta) = \delta(\Delta - \tau).$$

The mean is given by

$$\langle \Delta \rangle_{\text{bus}}^{\text{gap}} = \int_0^\infty \Delta \delta(\Delta - \tau) d\Delta = \tau.$$

For cars, we need $\frac{\Delta}{dt}$ chunks with no car followed by a single chunk with a car. Since the chunk with the car must be at the end of the sequence, there is no $n!$ term here. Thus,

$$\rho_{\text{car}}^{\text{gap}}(\Delta) dt = \lim_{dt \rightarrow 0} \left(1 - \frac{dt}{\tau}\right)^{\Delta/dt} \frac{dt}{\tau} = \frac{e^{-\Delta/\tau}}{\tau} dt.$$

The dt can be divided out. We can find the mean as well,

$$\langle \Delta \rangle_{\text{car}}^{\text{gap}} = \int_0^\infty \Delta \rho_{\text{car}}^{\text{gap}}(\Delta) d\Delta = \int_0^\infty \Delta e^{-\Delta/\tau} d\Delta / \tau = \tau.$$

(d) If another observer arrives at the road at a randomly chosen time, what is the probability distribution for the time Δ she has to wait for the first bus to arrive? What are the means of these distributions?

As noted in (b), the time until the next bus depends only on when the observer arrives, and is equally likely to be any time from 0 to τ . Thus, we have a uniform probability distribution,

$$\rho_{\text{bus}}^{\text{wait}}(\Delta) = \begin{cases} \frac{1}{\tau} & 0 \leq \Delta \leq \tau \\ 0 & \text{otherwise} \end{cases},$$

so that the mean is $\langle \Delta \rangle_{\text{bus}}^{\text{wait}} = \frac{\tau}{2}$.

Since the time until a car passes is completely independent of what happened before (no memory), we again conclude

$$\rho_{\text{car}}^{\text{wait}}(\Delta) = \frac{1}{\tau} e^{-\Delta/\tau},$$

with the mean again $\langle \Delta \rangle_{\text{car}}^{\text{wait}} = \tau$.

(e) In part (c), $\rho_{\text{car}}^{\text{gap}}(\Delta)$ was the probability that a randomly chosen gap was of length Δ . Write a formula for $\rho_{\text{car}}^{\text{time}}(\Delta)$, the probability that the second observer, arriving at a randomly chosen time, will be in a gap between cars of length Δ . From $\rho_{\text{car}}^{\text{time}}(\Delta)$, calculate the average length of the gaps between cars, using the time-weighted average measured by the second observer.

The probability distribution $\rho_{\text{car}}^{\text{time}}(\Delta)$ that a random time lies in a gap Δ can be written in terms of the probability distribution $\rho_{\text{car}}^{\text{gap}}(\Delta)$ that a random gap is of size Δ , by weighting each gap by the relative probability Δ that a random time falls inside that gap:

$$\rho_{\text{car}}^{\text{time}}(\Delta) = \Delta \rho_{\text{car}}^{\text{gap}}(\Delta) / \int \Delta \rho_{\text{car}}^{\text{gap}}(\Delta) d\Delta = \Delta e^{-\Delta/\tau} / \tau^2.$$

Alternatively, we can decompose the time Δ into the time t before the observer arrived and the time $\Delta - t$ after the observer arrived. If the gap is of length Δ then there must be some t for which a car passed at both of these times. Thus, we integrate over all the possible t ,

$$\rho_{\text{car}}^{\text{time}}(\Delta) = \int_0^\Delta \rho_{\text{car}}(t) \rho_{\text{car}}(\Delta - t) dt = \frac{\Delta}{\tau^2} e^{-\Delta/\tau},$$

where ρ_{car} is the result from part (d). Some may recognize this as a convolution $(\rho_{\text{car}} * \rho_{\text{car}})(\Delta)$. We see that this distribution is indeed normalized, and the mean is

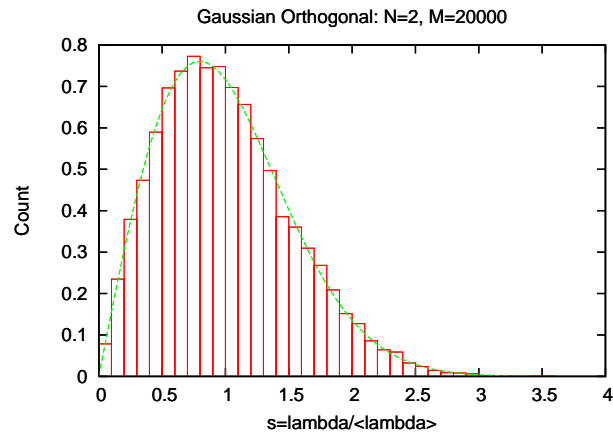
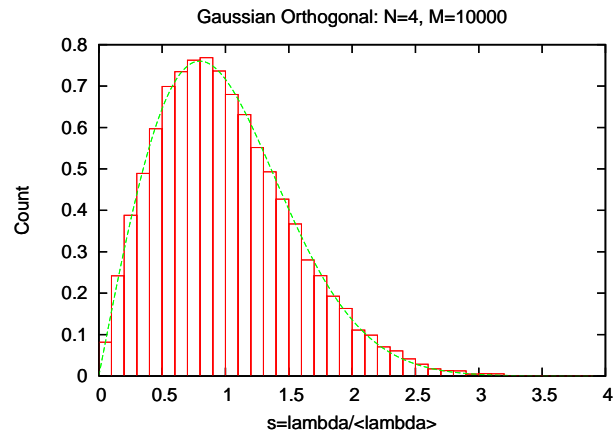
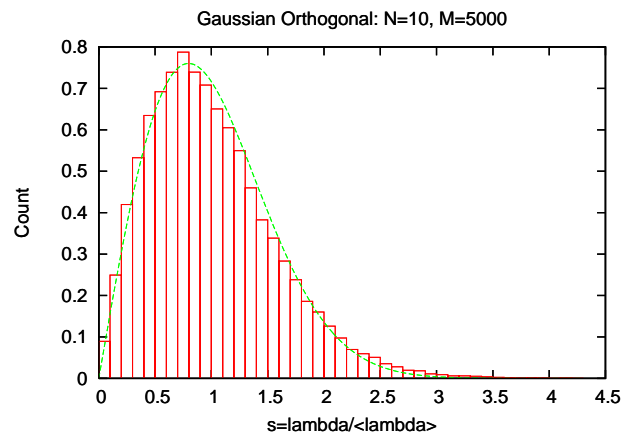
$$\langle \Delta \rangle_{\text{car}}^{\text{time}} = 2\tau.$$

1.6 Random Matrix Theory.

(a) Generate an ensemble with $M = 1000$ or so GOE matrices of size $N = 2, 4$, and 10 . Find the eigenvalues λ_n of each matrix, sorted in increasing order. Find the difference between neighboring eigenvalues $\lambda_{n+1} - \lambda_n$ for n , say, equal to $N/2$. Plot a histogram of these eigenvalue splittings divided by the mean splitting, with bin-size small enough to see some of the fluctuations.

See FIG. 1, 2, and 3.

(b) Show that the eigenvalue difference for M is $\lambda = \sqrt{(c-a)^2 + 4b^2} = 2\sqrt{d^2 + b^2}$ where $d = (c-a)/2$ and the trace is irrelevant. Ignoring the trace, the probability distribution of matrices can be written $\rho_M(d, b)$. What is the region in the (b, d) plane corresponding to the range of eigenvalues $(\lambda, \lambda + \Delta)$? If ρ_M is continuous and finite at $d = b = 0$, argue that the probability density $\rho(\lambda)$ of finding an eigenvalue splitting near $\lambda = 0$ vanishes (level repulsion). (Both d and b must vanish to make $\lambda = 0$.)

FIG. 1: Gaussian orthogonal, $N=2$ FIG. 2: Gaussian orthogonal, $N=4$ FIG. 3: Gaussian orthogonal, $N=10$

The eigenvalues are $(c + a)/2 \pm \sqrt{(c - a)^2 + 4b^2}/2$, so the eigenvalue difference is indeed $2\sqrt{d^2 + b^2}$. The region of the (b, d) plane corresponding to the range of eigenvalues considered is given by the annulus

$$\frac{\lambda^2}{4} \leq b^2 + d^2 \leq \frac{(\lambda + \Delta)^2}{4}$$

in the (b, d) plane with inner radius $\lambda/2$ and outer radius $(\lambda + \Delta)/2$. The area inside this annulus is $2\pi\lambda\Delta$ for small Δ , which vanishes for small eigenvalue splitting λ . Hence, so long as the probability density $\rho(d, b)$ of the ensemble is not singular at $d = b = 0$, the probability density for having a nearly degenerate eigenvalue pair separated by λ goes to zero proportionally to λ .

To get the two eigenvalues to agree, we need not only to have the two diagonal elements agree, but the off-diagonal element also to be zero. The probability density for this double-accident thus is zero.

(c) Calculate analytically the standard deviation of a diagonal and an off-diagonal element of the GOE ensemble. Calculate analytically the standard deviation of $d = (c - a)/2$ of the $N = 2$ GOE ensemble of part (b) and show that it equals the standard deviation of b .

For simplicity, consider a 2×2 matrix,

$$\begin{pmatrix} A & B \\ D & C \end{pmatrix},$$

where all the entries have standard deviation 1. Adding this to its transpose gives

$$M = \begin{pmatrix} a & b \\ b & c \end{pmatrix} = \begin{pmatrix} 2A & B + D \\ B + D & 2C \end{pmatrix},$$

so that $\sigma_a = 2\sigma_A = 2$, and likewise $\sigma_c = 2$. But $\sigma_b = \sqrt{\sigma_B^2 + \sigma_D^2} = \sqrt{2}$, and $\sigma_d = \frac{1}{2}\sqrt{\sigma_a^2 + \sigma_c^2} = \sqrt{2}$.

For larger GOE matrices $N > 2$ we can apply the same logic that diagonal elements are doubled while off-diagonal elements are added in quadrature, so that the standard deviations are 2 and $\sqrt{2}$, respectively.

(d) Calculate a formula for the probability distribution of eigenvalue spacings for the $N = 2$ GOE by integrating over the probability density $\rho_M(d, b)$.

We can now calculate

$$\rho(\lambda) = \int \rho_M(d, b) \delta(\lambda^2 - 4b^2 - 4d^2) dd db.$$

We know that d and b are independent Gaussians so that

$$\rho_M(d, b) = \frac{1}{4\pi} e^{-(b^2 + d^2)/4} = \frac{1}{4\pi} e^{-r^2/4},$$

where $r^2 = b^2 + d^2$. We then integrate $r dr d\phi$ instead of $dd db$. This brings out a 2π from the $\int d\phi$ and sets $\lambda = r/2$ from the δ -function so that

$$\rho(\lambda) = \frac{\lambda}{8} e^{-\lambda^2/16},$$

which is properly normalized because the conversion from r to λ has an extra factor of 2 come out in the denominator. Note that this is not a Gaussian and that $\rho(\lambda) = 0$ for $\lambda = 0$.

(e) Plot equation 1.6 along with your $N = 2$ results from part (a). Plot the Wigner surmise formula against $N = 4$ and $N = 10$ as well.

See the figures referenced in (a)

(f) Generate an ensemble with $M = 1000 \pm 1$ symmetric matrices with size $N = 2, 4$, and 10 . Plot the eigenvalue distributions as in part (a). Are they universal for $N = 2$ and 4 ? Do they appear to be nearly universal for $N = 10$? Plot the Wigner surmise along with your histogram for $N = 10$.

See FIG. 4, 5, and 6. For small matrix size N the behavior is clearly different from that of the GOE ensemble, but by $N = 10$ the agreement is excellent.

(g) Show that $\text{Tr}[H^T H]$ is the sum of the squares of all elements of H . Show that this trace is invariant under orthogonal coordinate transformations.

Consider $\text{Tr}[H^T H] = \sum_i [H^T H]_{ii}$. But we can expand the matrix product $[H^T H]_{ii} = \sum_j H_{ij}^T H_{ji}$:

$$\text{Tr}[H^T H] = \sum_{ij} H_{ij}^T H_{ji} = \sum_{ij} H_{ji} H_{ji} = \sum_{ij} (H_{ji})^2.$$

So we see that $\text{Tr}[H^T H]$ is the sum of the squares of all the elements of H .

Now define $M = R^T H R$ to be an orthogonal transformation of H . We find that

$$\text{Tr}[M^T M] = \text{Tr}[(R^T H R)^T (R^T H R)] = \text{Tr}[R^T H^T R R^T H R] = \text{Tr}[R^T H^T H R] = \text{Tr}[H^T H R R^T] = \text{Tr}[H^T H],$$

where we use the cyclic invariance of trace and the condition $R^T R = R R^T = 1$.

(h) Write the probability density $\rho(H)$ for finding GOE ensemble member H in terms of the trace formula in part (g). Argue, using your formula and the invariance from part (g), that the GOE ensemble is invariant under orthogonal transformations: $\rho(R^T H R) = \rho(H)$.

If H is an N by N member of GOE then it has $N(N+1)/2$ independent elements (the diagonal and half of the off-diagonal elements). The diagonal elements each have standard deviation 2, while the off-diagonals have a standard deviation of $\sqrt{2}$. Thus the probability density of H is

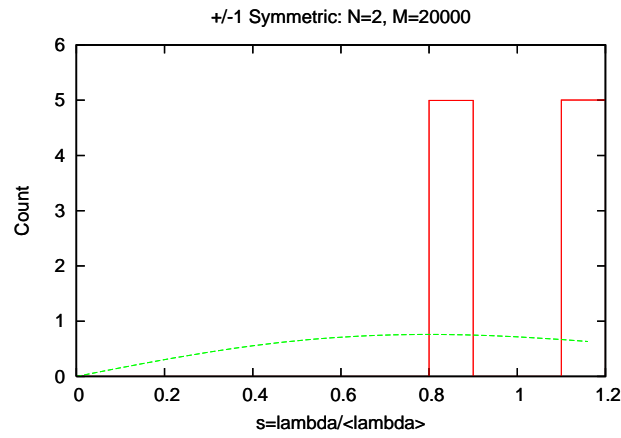
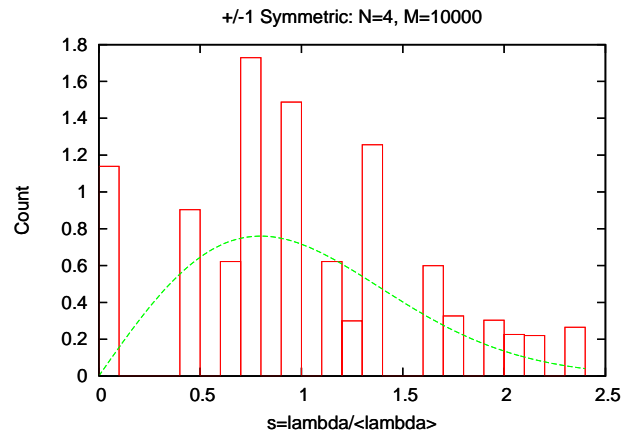
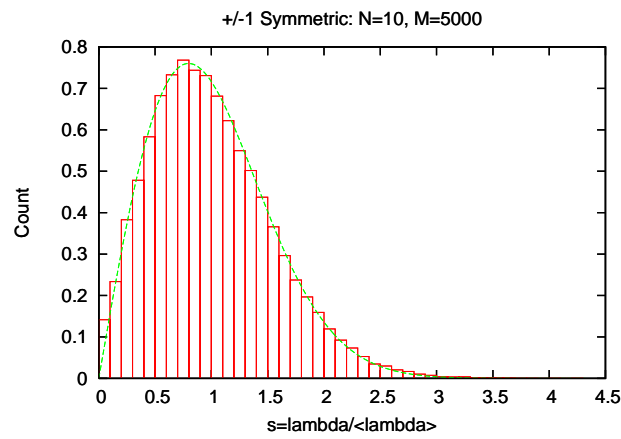
$$\rho(H) = \prod_{i \leq j} \rho(H_{ij}) = \left(\prod_i \rho(H_{ii}) \right) \left(\prod_{i < j} \rho(H_{ij}) \right).$$

We substitute the proper Gaussian distributions for $\rho(H_{ij})$, change the second product to double count, and add a radical to correct this:

$$\rho(H) = \left(\prod_i \frac{1}{\sqrt{8\pi}} e^{-H_{ii}^2/8} \right) \sqrt{\prod_{i \neq j} \frac{1}{\sqrt{4\pi}} e^{-H_{ij}^2/4}} = \left(\frac{1}{\sqrt{8\pi}} \right)^N \left(\frac{1}{\sqrt{4\pi}} \right)^{N(N-1)/2} \prod_{i,j} e^{-H_{ij}^2/8}.$$

Finally, we collect the product into a sum in the exponent which equals the trace, so that

$$\rho(H) = \left(\frac{1}{2} \right)^{N(N+2)/2} \left(\frac{1}{\sqrt{\pi}} \right)^{N(N+1)/2} \exp \left(-\frac{1}{8} \text{Tr}[H^T H] \right).$$

FIG. 4: ± 1 symmetric, $N=2$ FIG. 5: ± 1 symmetric, $N=4$ FIG. 6: ± 1 symmetric, $N=10$

We see now that, since $\text{Tr}[H^T H]$ is invariant under orthogonal transforms, $\rho(H)$ is necessarily invariant as well, since it depends only on this trace.

2.1 Random Walks in Grade Space.

(a) What is the expected mean and standard deviation for the exam?

Let's define g_i as the grade for problem i , with $1 \leq i \leq N = 10$. We thus have $P(g_i = 10) = 0.7$, $P(g_i = 0) = 0.3$, $\langle g_i \rangle = 7$, $\langle g_i^2 \rangle = 70$ and $\sigma_{g_i} = \sqrt{21}$.

Next, we can define the total grade on the exam $G = \sum_{i=1}^N g_i$. Then, we have:

$$\begin{aligned} \langle G \rangle &= \sum_{i=1}^N \langle g_i \rangle = N \langle g_i \rangle \\ &= 70 \\ \sigma_G &= \sqrt{\langle G^2 \rangle - \langle G \rangle^2} = \sqrt{\sum_{i,j=1}^N (\langle g_i g_j \rangle - \langle g_i \rangle \langle g_j \rangle)} \\ &= \sqrt{\sum_{i=1}^N (\langle g_i^2 \rangle - \langle g_i \rangle^2)} = \sqrt{N} \sigma_{g_i} \\ &\simeq 14.5 \end{aligned}$$

(b) What physical interpretation do you make of the ratio of the random standard deviation and the observed one?

The ratio is very close to 1. Multiple-choice tests with a few heavily-weighted questions are often unpopular, as students feel that their scores are often as much luck as they are testing their knowledge. This exercise quantifies that feeling; the random statistical fluctuations in ten multiple-choice questions is roughly as large as the total range of performance expected in a typical (differently graded) exam. If this one exam were the only grade in a course, luck and skill would be equally weighted. If there are several ten-question exams, the statistical fluctuations will tend to average out and the differences due to skill will become more evident.

2.2 Photon diffusion in the Sun.

About how many random steps N will the photon take of length ℓ to get to the radius R where convection becomes important? About how many years δt will it take for the photon to get there?

We know for random walks that $\langle R \rangle \sim \ell \sqrt{N}$ so that

$$N \approx \left(\frac{R}{\ell} \right)^2 \approx 10^{26},$$

where we want a radius $R = 5 \times 10^8 \text{m}$ and we have a mean free path $\ell = 5 \times 10^{-5} \text{m}$. Such a mean free path gives a scattering time $\tau = \frac{\ell}{c} \approx 1.7 \times 10^{-13} \text{s}$ so that N steps will take

$$T \approx N\tau \approx 1.7 \times 10^{13} \text{s} \approx 5 \times 10^5 \text{yr}.$$

2.5 Generating Random Walks.

(a) Generate 1 and 2 dimensional random walks.

See FIG. 7, 8, 9, 10. Notice that the scale for the random walks grows approximately as \sqrt{N} , so the 1000 step walks span ± 10 where the 10 step walks span ± 1 .

(b) Generate a scatter plot for 10000 2d random walks with 1 step and 10 steps.

See FIG. 11. Note the emergent spherical symmetry.

(c) Calculate the RMS stepsize a for a one-dimensional random walk. Compare central limit theorem result to histograms.

The stepsize can be calculated simply:

$$a = \sqrt{\langle (\Delta x)^2 \rangle} = \sqrt{\int_{-\frac{1}{2}}^{\frac{1}{2}} dx x^2} = \frac{1}{2\sqrt{3}} = 0.289$$

Thus the variance should be given by $\sigma = \frac{\sqrt{N}}{2\sqrt{3}}$.

See FIG. 12, 13, 14, 15. The distribution is triangular for $N = 2$ steps, but remarkably Gaussian for $N > 3$.

2.6 Fourier and Green.

(a)

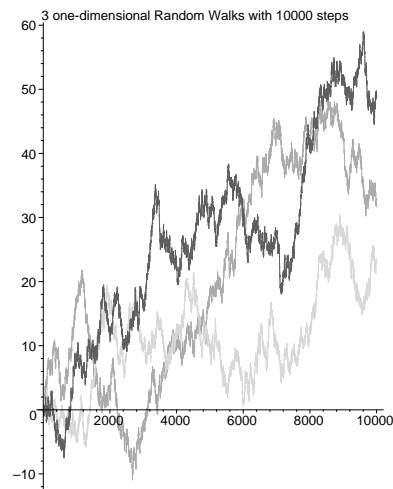
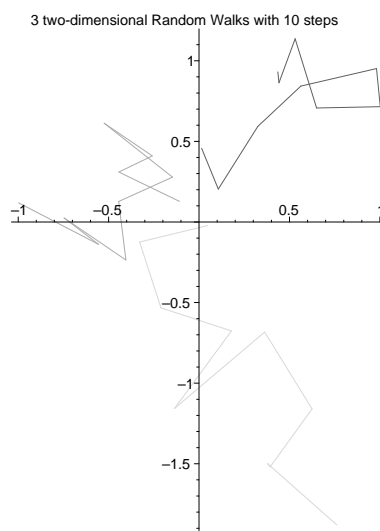
$\rho(x, 0) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, Fourier transforms to $\tilde{\rho}_k(0) = e^{-k^2/2}$. This will time evolve as $\tilde{\rho}_k(t) = \tilde{\rho}_k(0) e^{-Dk^2 t}$, so that the effect on the Gaussian is to simply increase the spread while decreasing the amplitude: $\tilde{\rho}_k(t) = e^{-\frac{k^2}{2}(\sigma^2 + 2Dt)}$, where $\sigma = 1\text{m}$. With $D = 0.001\text{m}^2/\text{s}$ and $t = 10\text{s}$ we see a 2% change in σ^2 , which is a 1% change in the width and a 1% attenuation in the height.

(b)

For the sinusoid, we have a single important Fourier component: $k = 10/\text{m}$. Now we again have $\tilde{\rho}_k(t) = \tilde{\rho}_k(0) e^{-Dk^2 t}$, but at $k = 10/\text{m}$ we now have a factor $1/e$ which is an attenuation of 63%.

(c) **Pictures.** Now consider time evolution for the next ten seconds. The initial density profile $\rho(x, t = 0)$ is again shown at left. Which of the choices in figure 2.11 represents the density at $t = 10\text{s}$?

Thus, the sinusoid should be much smaller than the Gaussian and the correct choice is ☒ C.

FIG. 7: 1 dimensional random walks $x(t)$ for 10000 stepsFIG. 8: 2 dimensional random walks (x,y) for 10 steps

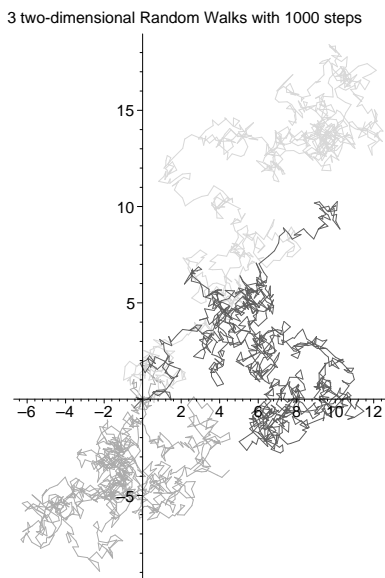


FIG. 9: 2 dimensional random walks (x,y) for 1000 steps

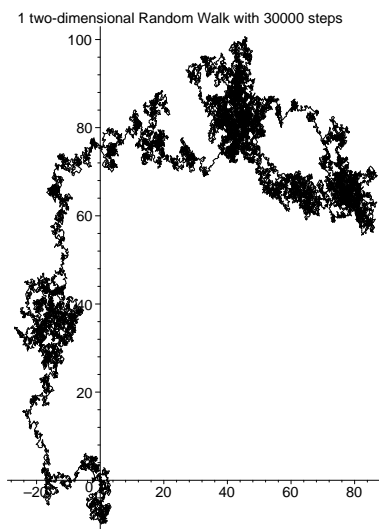


FIG. 10: 2 dimensional random walks (x,y) for 30000 steps

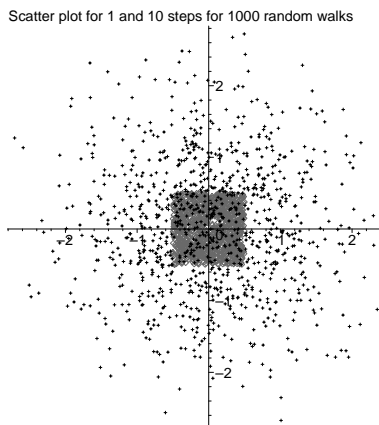


FIG. 11: Scatter plot for 1000 random walks with step 1 and step 10

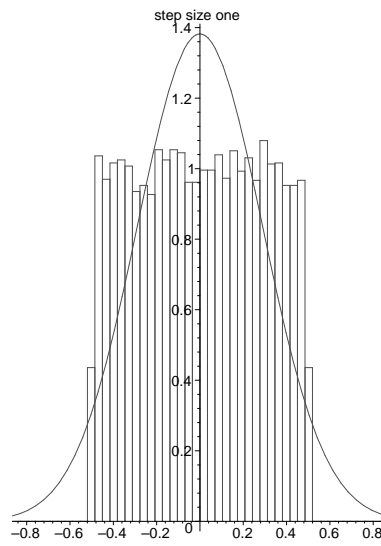


FIG. 12: Histogram representing the end points of 10000 1-step random walks. Histogram is renormalized to have area 1.

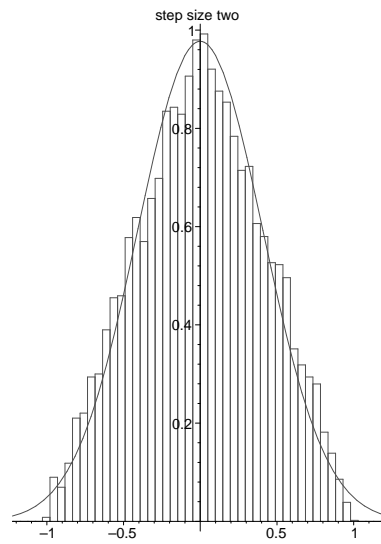


FIG. 13: Histogram representing the end points of 10000 2-steps random walks. Histogram is renormalized to have area 1.

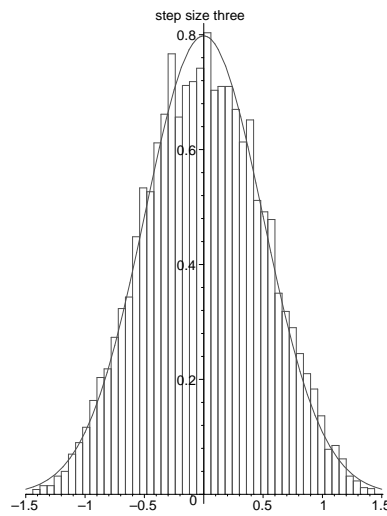


FIG. 14: Histogram representing the end points of 10000 3-steps random walks. Histogram is renormalized to have area 1.

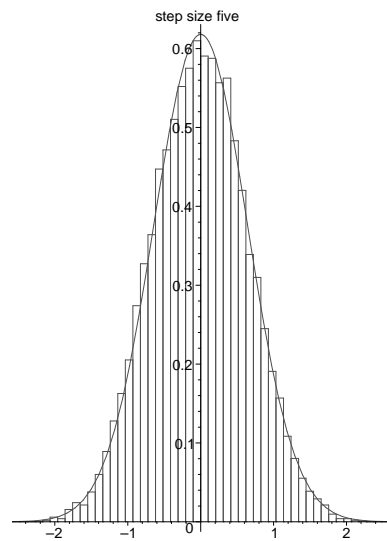


FIG. 15: Histogram representing the end points of 10000 5-steps random walks. Histogram is renormalized to have area 1.

2.8 Polymers and Random Walks.

(a) If the persistence length to bending DNA is 50nm with 3.4Å per nucleotide base pair, what will the root-mean-square distance $\sqrt{\langle R^2 \rangle}$ be between the ends of a gene in solution with 100,000 base pairs, if the DNA is accurately represented as a random walk?

If the DNA is a self-intersecting random walk, then we expect a \sqrt{N} dependence on the number of steps. 100,000 base pairs has a total length of 34,000nm, which is $N = 680$ persistence lengths. If we take this to be the step size then we expect a radius of $\sqrt{N} \sim 26$ persistence lengths, or

$$\sqrt{\langle R^2 \rangle} = 1300\text{nm}.$$

(b) Two dimensional self-avoiding random walk

Give a convincing, short argument explaining whether or not a typical, non-self-avoiding random walk in two dimensions will come back after a large number of monomers and cross itself.

There are a number of arguments one could think of. Following is an argument very much in the spirit of statistical mechanics (particularly mean field theory). We know that a typical non-self-avoiding random walk of N (unit) steps will roughly fill a ball of radius \sqrt{N} . The volume of this ball in d dimensions scales as $N^{d/2}$. We consider the walk to live on a lattice. Now the probability that a given pair monomers lie on the same lattice site is $\frac{1}{N^{d/2}}$. But with N monomers, there are N^2 different pairs, so the probability that any pair will intersect is $N^{2-d/2}$. As $N \rightarrow \infty$, this probability diverges for $d \leq 4$, and the walk will most likely cross itself. For $d > 4$, it vanishes and self-intersection is no longer an issue.

(c) BU java applet

Measure for a reasonable length of time, print out the current state, and enclose it. Did the simulations give $R \sim \sqrt{L}$? If not, what's the estimate that your simulation gives for the exponent relating R to L ? How does it compare with the two-dimensional theoretical exponent given on the Web site?

The applet should give an exponent around 3/4, in agreement with theory.

2.11 Stocks, Volatility and Diversification .

If the stock market is a random walk, then l is the number of steps in the random walk and P_l is the final position of the walk. Defining a walk for all t , ie $P_l(t)$, gives us an ensemble of walks with which to do statistics. Of course, for our results on random walks to apply, the random walks labeled by different t 's should be independent!

(a) Plot the price index versus time.

See FIG. 16: There was no especially dramatic drop after 9/11 (day 6903) — just a small apparent extra dip in an overall decline that started several months earlier. Most of the fluctuation in stock prices appears to be internal, not triggered by external events.

(b) Plot histograms of changes for various time steps. How long should you stay in the market? Why are the yearly changes so different?

See FIG. 17 18 19: Clearly, one should stay in the market for years to make money. The other distributions look Gaussian with means that are small compared to the fluctuations: you're nearly as likely to lose money as earn it for times smaller than a year.

The yearly changes are not Gaussian distributed because the walks in our ensemble are not independent. Although there are many intervals of 200 trading days in our sample, the intervals overlap a great deal; there are only a few non-overlapping regions (one per year).

(c) Show that the logarithm of a gaussian is an inverted parabola. Plot the logarithm of the histogram for the weekly percentage changes from part b). Are there more large percentage changes than expected from a Gaussian distribution or fewer?

For a density of $\rho(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{x^2}{\sigma^2}}$, we have $\ln(\rho(x)) = -\frac{1}{2}\ln(2\pi\sigma^2) - \frac{x^2}{2\sigma^2}$, which is an inverted parabola. As can be seen on FIG. 20, our distribution seems to have a "fat tail".

(d) Plot volatility; does it behave as it should?

We have $v_1 = 1.06$, $v_5 = 2.29$ and $v_{20} = 4.30$.

See FIG. 21: Random walk theory predicts that the variance of the probability density of the final position of a N -step walk goes as \sqrt{N} . In the language of volatility, this means that $v_l \simeq \sqrt{l}$; as can be seen on the figure, this is clearly the case.

e) Suppose these five hundred stocks had mean annual percentage returns m_i and each had mean volatility σ_i . Suppose they were equally weighted in the index, and their fluctuations were uncorrelated. What would the return and volatility for the index be? Without inside information or insight as to which stocks will have higher mean returns, is there any average disadvantage of buying portions of each stock over buying the index? Which has lower volatility?

To answer this question fully, let's define a few things. Let x_i be a random variable representing the percentage return of stock $i \leq N = 500$. Then $\langle x_i \rangle = m_i$ and $\langle (x_i - m_i)^2 \rangle = \sigma_i^2$. We'll also define a "strategy" $\{S_i\}$ such that S_i represents the fraction of our portfolio composed by stock i . Thus:

$$\sum_{i=1}^N S_i = 1$$

Finally, we'll define x_S as the percentage return of our portfolio when we use strategy $\{S_i\}$. Hence:

$$x_S = \sum_{i=1}^N S_i x_i \Rightarrow m_S = \langle x_S \rangle = \sum_{i=1}^N S_i m_i$$

"Buying the index" means opting for a strategy where all stocks are weighted in equally, ie $S_i = \frac{1}{N} \forall i$. Hence:

$$m_{\text{index}} = \frac{1}{N} \sum_{i=1}^N m_i$$

Without insider information, it is impossible to choose an optimal strategy! Thus if we choose a strategy at random, then on average (average over strategies!) the mean return will be:

$$m_{\langle S \rangle} = \sum_{i=1}^N \langle S_i \rangle_{\langle S \rangle} m_i = \frac{1}{N} \sum_{i=1}^N m_i$$

because the probability of choosing a given strategy is uniformly distributed and thus $\langle S_i \rangle = \langle S_1 \rangle = \frac{1}{N} \forall i$. Thus we see that choosing our own strategy at random or buying the index yields the same average return! Volatility, however, is another story. The volatility of a given strategy S is given by:

$$\sigma_S = \sqrt{\langle (x_S - m_S)^2 \rangle} = \dots = \sqrt{\sum_{i=1}^N S_i^2 \sigma_i^2}$$

Hence $\sigma_{\text{index}} = \frac{1}{N} \sqrt{\sum_{i=1}^N \sigma_i^2}$. For somebody picking a strategy at random, the average volatility (average over strategies!) is given by:

$$\sigma_{\langle S \rangle} = \langle \sqrt{\sum_{i=1}^N S_i^2 \sigma_i^2} \rangle_{\langle S \rangle}$$

This I don't know how to evaluate exactly, but we can get some intuition by considering only the subset of strategies where only one stock is chosen (let's call this set S'). Then, it is easy to see that $\sigma_{S'} = \frac{1}{N} \sum_{i=1}^N \sigma_i$, which is larger than σ_{index} . Thus, it seems reasonable to think that the volatility of the index is smaller than the average volatility of a randomly chosen strategy.

2.12 Computational Finance: Pricing Derivatives.

(a) What value of α makes the two possible final values of the portfolio equal? What is the common final value P_F ?

Let $\hat{\alpha}$ be the value of α that make both final values equal. Then

$$\hat{\alpha} = \frac{V_d - V_u}{X_u - X_d}$$

$$P_F = \frac{V_d X_u - X_d V_u}{X_u - X_d}$$

(b) What initial value of V_0 of the derivative makes the initial value of the portfolio equal to the final value?

We'll assume both final values of the portfolio are equal:

$$P_0 = V_0 + \hat{\alpha} X_0 = P_F$$

$$V_0 = P_F - \hat{\alpha} X_0$$

$$V_0 = \frac{V_d(X_u - X_0) - V_u(X_d - X_0)}{X_u - X_d}$$

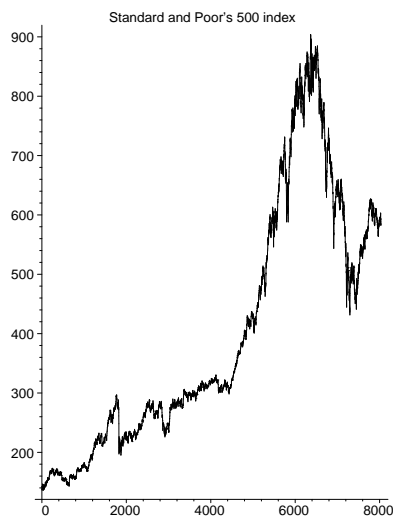


FIG. 16: Standard and Poor's 500 index as a function of time, in days, since 1982. The world trade center attack (day 6903) did not trigger the overall downturn on this figure.

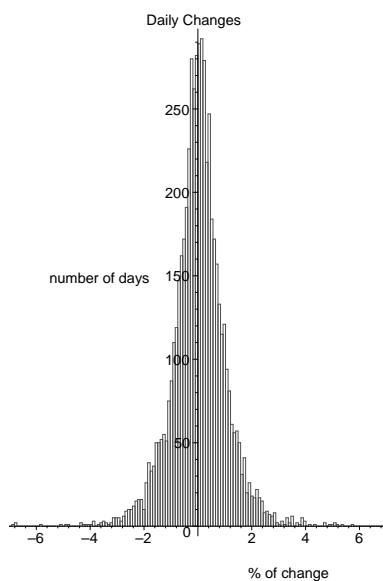


FIG. 17: Histogram of daily changes.

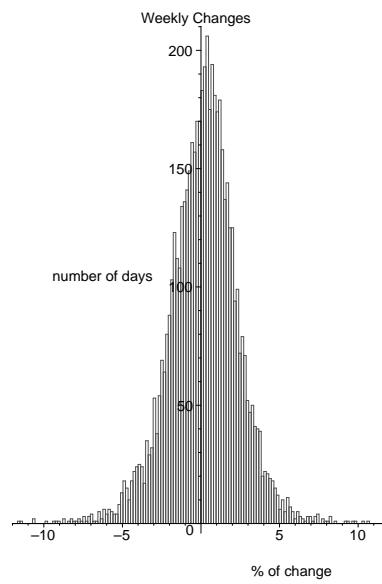


FIG. 18: Histogram of weekly changes.

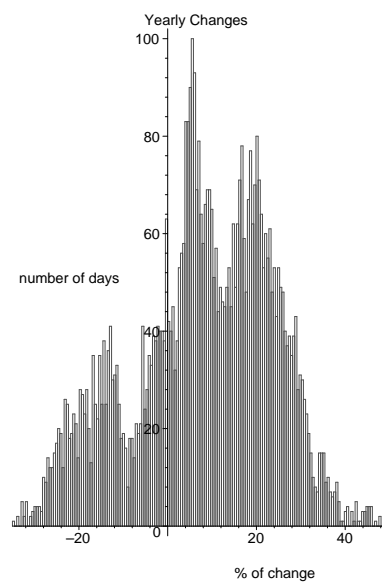


FIG. 19: Histogram of yearly changes.

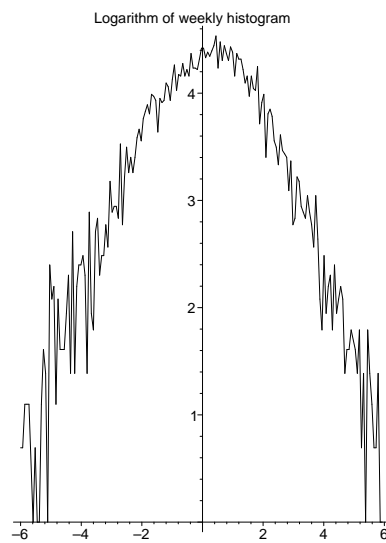
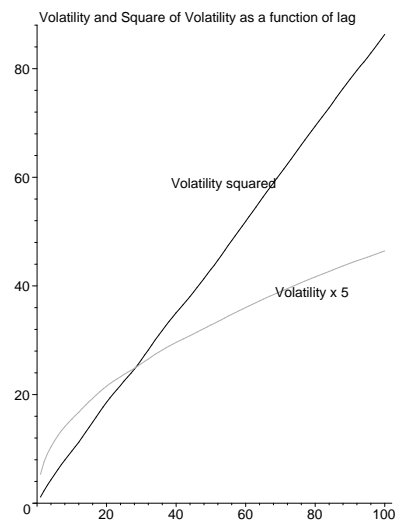


FIG. 20: logarithm of Histogram of weekly changes.

FIG. 21: $5 \times v_l$ and v_l^2 for lag between 1 and 100 days.

(c) Does your answer depend upon the probabilities p_u and p_d of going up or down?

No! Both situations lead to the same value for the portfolio, P_F . Thus, the final value of the portfolio is P_F with probability one.

3.5 Hard Sphere Gas.

(a) The area allowed for the second disc is $A - \pi(2r)^2$, ignoring the small correction when the excluded region around the first disc overlaps the excluded region near the walls of the box. What is the allowed $2N$ -dimensional volume in configuration space, of allowed zero-energy configurations of hard discs, in this dilute limit?

We generalize the $N = 2$ result $A_2 = (A)(A - \pi(2r)^2)$ by realizing that each successive disc has $d = \pi(2r)^2$ less area than the previous. Thus, we multiply all these areas together, finally adding a factor of $1/N!$ because there are $N!$ equivalent permutations of placing N discs in a configuration. Thus,

$$A_N = \frac{1}{N!} \prod_{n=1}^N (A - (n-1)d).$$

(b) What is the configurational entropy for the hard discs? Here, simplify your answer so that it does not involve a sum over N terms, but valid to first order in the area of the discs πr^2 . Show, for large N , that it is well approximated by $S_Q = Nk_B(1 + \log(A/N - b))$, with b representing the effective excluded area due to the other discs. What is the value of b , in terms of the area of the disc?

The configurational entropy is given by

$$S_Q = k_B \log A_N = k_B \left[-\log N! + \sum_{n=1}^N \log (A - (n-1)d) \right].$$

But we are asked to write this without a sum. Thus, we expand in the dilute limit that $Nd \ll A$, so that $\log(A - [n-1]d) = \log A + \log(1 - [n-1]d/A) = \log A - [n-1]\epsilon$, where $\epsilon = d/A$, and we have used the Taylor series $\log(1+\epsilon) = \epsilon + \mathcal{O}(\epsilon^2)$. Now that the n is outside the log, we can easily sum this arithmetic series, $\sum_{n=1}^N (n-1) = N(N-1)/2$, and use Stirling's approximation $\log(N!) \sim N \log N - N$,

$$S_Q/k_B = -(N \log N - N) + \left(N \log A - \frac{N(N-1)}{2} \epsilon \right).$$

We divide out one of the N 's and recombine the log to find

$$S_Q/Nk_B = 1 + \log \left(\frac{A}{N} - \frac{A\epsilon}{N} \frac{N-1}{2} \right) = 1 + \log \left(\frac{A}{N} - b \right),$$

where

$$b = \frac{A\epsilon}{N} \frac{N-1}{2} \approx 2\pi r^2$$

is twice the area of the discs, or half the excluded area d from each disc.

(c) Find the pressure for the hard-sphere gas in the large N approximation of part (b). Does it reduce to the ideal gas law for $b = 0$?

Pressure is the work done to expand by a unit volume (area), so to find the pressure we must see what happens when we change the area. In three dimensions, eqns. 3.31 and 3.47 tell us that $\partial S / \partial V|_E = P/T$; the analogy in two dimensions here is

$$P = T \left. \frac{\partial S}{\partial A} \right|_E.$$

But the energy here is purely a kinetic energy, which is $Nk_B T$ ($\frac{1}{2}k_B T$ for $2N$ velocity coordinates), so keeping a constant energy amounts to keeping a constant temperature. Hence the momentum components of the entropy do not change as we change the area at constant energy. We can hence substitute our expression for S (the configurational entropy) and differentiate to get

$$P(A - bN) = Nk_B T,$$

which is familiarly somewhere between the ideal gas law and the van der Waals law. When $b = 0$, this reduces to the even more familiar $PA = nk_B T$.

3.6 Connecting Two Macroscopic Systems.

Derive formula 3.23 for the volume of the energy surface of the whole system using δ -functions.

We begin with expressions for the volumes of the energy surfaces in terms of δ -functions:

$$\Omega_1(E_1) = \int d\mathbb{Q}_1 d\mathbb{P}_1 \delta(E_1 - \mathcal{H}_1(\mathbb{Q}_1, \mathbb{P}_1))$$

$$\Omega_2(E_2) = \int d\mathbb{Q}_2 d\mathbb{P}_2 \delta(E_2 - \mathcal{H}_2(\mathbb{Q}_2, \mathbb{P}_2))$$

$$\Omega(E) = \int d\mathbb{Q}_1 d\mathbb{P}_1 d\mathbb{Q}_2 d\mathbb{P}_2 \delta(E - \mathcal{H}_1 - \mathcal{H}_2).$$

The definition of δ -functions gives us

$$\int dE_1 \delta(E_1 - \mathcal{H}_1) = 1,$$

for any $\mathbb{Q}_1, \mathbb{P}_1$. We insert this into the $\Omega(E)$ integral.

$$\Omega(E) = \int d\mathbb{Q}_1 d\mathbb{P}_1 d\mathbb{Q}_2 d\mathbb{P}_2 dE_1 \delta(E - \mathcal{H}_1 - \mathcal{H}_2) \delta(E_1 - \mathcal{H}_1).$$

We may now replace $\mathcal{H}_1 \rightarrow E_1$ while retaining the δ -function. Rearranging the integrals, we can decouple the systems:

$$\Omega(E) = \int dE_1 \int d\mathbb{Q}_1 d\mathbb{P}_1 \delta(E_1 - \mathcal{H}_1) \int d\mathbb{Q}_2 d\mathbb{P}_2 \delta(E - E_1 - \mathcal{H}_2).$$

But these two integrals are just $\Omega_1(E_1)$ and $\Omega_2(E - E_1)$. Thus,

$$\Omega(E) = \int dE_1 \Omega_1(E_1) \Omega_2(E - E_1).$$

3.8 Microcanonical Energy Fluctuations.

(a) Show that $\frac{1}{k_B} \frac{\partial^2 S}{\partial E^2} = -\frac{1}{k_B T} \frac{1}{N c_v T}$

This is just a derivative juggling exercise:

$$\frac{1}{k_B} \frac{\partial^2 S}{\partial E^2} = \frac{1}{k_B} \frac{\partial}{\partial E} \left(\frac{1}{T} \right) = \frac{1}{k_B} \frac{\partial T}{\partial E} \frac{\partial}{\partial T} \left(\frac{1}{T} \right) = -\frac{1}{k_B T} \frac{1}{N c_v T}$$

(b) Show equation (3.64).

Assume the two subsystems have the same number of particles N , and are at the same temperature T . Also, we'll now call E_1 simply E . Then, from eqn (3.63):

$$\frac{1}{\sigma_E^2} = -\frac{1}{k_B} \frac{\partial^2 S_1}{\partial E^2} - \frac{1}{k_B} \frac{\partial^2 S_2}{\partial E_2^2}.$$

Using the result of part (a), this implies:

$$\frac{1}{\sigma_E^2} = \frac{1}{k_B N T^2 c_v^{(1)}} + \frac{1}{k_B N T^2 c_v^{(2)}},$$

hence the final result:

$$\frac{1}{c_v^{(1)}} + \frac{1}{c_v^{(2)}} = \frac{k_B N T^2}{\sigma_E^2}$$

(c) Solve for the total specific heat of the molecular dynamics simulation:

First, we note that $\sigma_K = \sigma_{E_1}$; in a microcanonical simulation the total energy (potential + kinetic) is conserved, so the fluctuations in one subsystem must be exactly minus the fluctuations in the other. Solving eqn (3.65) for $c_v^{(1)}$ using our formula for $c_v^{(2)} = (3/2)k_B$ and solving $K = 3/2 N k_B T$ to get $T = (2/3)K/N k_B$, we get $c_v^{(1)} = 9k_B N \sigma_K^2 / (4K^2 - 6N \sigma_K^2)$. Then the total specific heat per particle is

$$c_v = c_v^{(1)} + c_v^{(2)} = k_B \frac{K^2}{\frac{2}{3}K^2 - N \sigma_K^2}$$

3.9 Gauss and Poisson.

(a) Find the exact formula for this probability: n particles in V , with a total of T particles in KV .

We consider the configuration in which the first n particles are in the subvolume V and the final $T - n$ particles are outside this subvolume (and thus occupy a volume $(K - 1)V$). The total available volume of this configuration is then

$$V^n ([K - 1]V)^{T-n} = V^T [K - 1]^{T-n}.$$

The total volume of all configurations is simply $(KV)^T$. We divide to find the probability of exactly the first n particles residing in the subvolume. But there are $\binom{T}{n}$ different ways to pick these n particles so that the total probability is

$$P(n) = \frac{[K - 1]^{T-n}}{K^T} \binom{T}{n} = \left(\frac{K - 1}{K} \right)^{T-n} \frac{1}{K^n} \frac{T!}{n!(T - n)!}.$$

(b) Show that the Poisson distribution is normalized: $\sum_n \rho_n = 1$. Calculate the mean of the distribution $\langle n \rangle$ in terms of a . Calculate the standard deviation $\langle (n - \langle n \rangle)^2 \rangle$.

First we check normalization:

$$\sum_{n=0}^{\infty} \rho_n = e^{-a} \sum_{n=0}^{\infty} \frac{a^n}{n!} = e^{-a} \cdot e^a = 1,$$

where we recognized $\sum \frac{a^n}{n!}$ as the Taylor series for e^a .

The mean of the Poisson distribution is given by

$$\langle n \rangle = \sum_{n=0}^{\infty} n \frac{a^n}{n!} e^{-a}.$$

We note that the $n = 0$ term vanishes, so we remove it from the sum. Now we absorb $\frac{n}{n!} = \frac{1}{(n-1)!}$ and pull an a out of the power, $a^n = aa^{n-1}$. Now we can reindex $n \rightarrow n - 1$ leaving

$$\langle n \rangle = \sum_{n=1}^{\infty} \frac{aa^{n-1}}{(n-1)!} e^{-a} = a \sum_{n=0}^{\infty} \frac{a^n e^{-a}}{n!} = a \sum_{n=0}^{\infty} \rho_n = a,$$

since this is just the normalization condition.

We treat the mean square similarly, doing the same process twice:

$$\langle n^2 \rangle = \sum_{n=0}^{\infty} n^2 \frac{a^n}{n!} e^{-a} = a \sum_{n=0}^{\infty} (n+1) \rho_n = a \langle n+1 \rangle = a \langle n \rangle + a = a^2 + a.$$

Thus the standard deviation is

$$\langle (n - \langle n \rangle)^2 \rangle = \langle n^2 \rangle - \langle n \rangle^2 = a.$$

(c) As $K \rightarrow \infty$, show that the probability that n particles fall in the subvolume V has the Poisson distribution 3.64. What is a ?

We begin with the solution to (a),

$$P(n) = \frac{[K-1]^{T-n}}{K^T} \binom{T}{n} = \left(1 - \frac{1}{K}\right)^{T-n} \frac{1}{K^n} \frac{T!}{n!(T-n)!}.$$

In the limit $K \rightarrow \infty$ and $n \ll T \rightarrow \infty$, we can take $(1 - 1/K)^{T-n} = \exp(T/K)$. Additionally we use Stirling's approximation, $T! \approx (T/e)^T$, to simplify

$$\frac{T!}{(T-n)!} \approx \frac{(T/e)^T}{((T-n)/e)^{(T-n)}} = \left(\frac{T-n}{T}\right)^{n-T} \left(\frac{T}{e}\right)^n \approx e^{\frac{n}{T}(T-n)} \left(\frac{T}{e}\right)^n \approx T^n.$$

This gives

$$P(n) = e^{-T/K} \frac{(T/K)^n}{n!},$$

which is a Poisson distribution with mean $a = \frac{T}{K}$.

(d) Using your result from part (a), write the exact formula for $\log(P_m)$. Apply the weak Stirling's formula. Expand your result around $m = 0$ to second order in m , and show that $\log(P_m) \approx -m^2/2\sigma_K^2$, giving a Gaussian form $P_m \sim \exp(-m^2/2\sigma_K^2)$. What is σ_K ? In particular, what is σ_2 and σ_∞ ?

We define $N_0 = \frac{T}{K}$ and $m = n - N_0$, with $|m| \ll N_0$. We will leave P_m in terms of n for now.

$$P_m = \frac{T!}{n!(T-n)!} \left(1 - \frac{1}{K}\right)^T (K-1)^{-n}$$

so that a log, followed by the weak form of Stirling's approximation, gives

$$\begin{aligned} \log P_m &= T \log T - T - n \log n + n - (T-n) \log(T-n) + (T-n) + T \log \left(1 - \frac{1}{K}\right) - n \log(K-1) \\ &= T \log T - n \log n - (T-n) \log(T-n) + T \log \left(1 - \frac{1}{K}\right) - n \log(K-1). \end{aligned}$$

We differentiate $\partial/\partial m = \partial/\partial n$,

$$\begin{aligned} \frac{\partial}{\partial n} \log P_m &= -1 - \log n + 1 + \log(T-n) - \log(K-1) \\ \frac{\partial^2}{\partial n^2} \log P_m &= -\frac{1}{n} - \frac{1}{T-n}. \end{aligned}$$

We see that the first derivative vanishes at $m = 0$ ($n = N_0$) and the second derivative becomes

$$(\log P_m)''(0) = -\frac{1}{N_0} \frac{K}{K-1} = -\frac{1}{\sigma_K^2}.$$

Therefore we have

$$P_m \sim e^{-\frac{m^2}{2N_0} \frac{K-1}{K}}.$$

When $K = 2$ we have $\sigma_2^2 = T/2$ which agrees with the result from Section 3.2.1. When $K = \infty$, we have $\sigma_\infty^2 = N_0$, which agrees with eqn 3.67.

(e) In terms of T and n , what factor would the square-root term have contributed if you had kept it in Stirling's formula going from part (a) to part (d)? Show from equation 3.66 that the fluctuations are small, $m = n - N_0 \ll N_0$ for large N_0 . Ignoring these fluctuations, set $n = N_0$ in your factor, and give the prefactor multiplying the Gaussian in equation 3.66.

We can work through the previous part with the strong form of Stirling's approximation, resulting in an extra factor of $\sqrt{2\pi T}/\sqrt{2\pi n}\sqrt{2\pi(T-n)}$. This factor sticks around through the entire derivation and leaves a properly normalized distribution

$$P_m \approx \frac{1}{\sqrt{2\pi N_0}} \sqrt{\frac{K}{K-1}}.$$

We can calculate the fluctuations to be

$$\frac{\sigma}{N_0} \sim N_0^{-1/2},$$

which is very small for large N_0 .

3.10 Triple Product Relation.

(a) Show by solving equation 3.75 for dS that $\frac{\partial S}{\partial V}|_{N,E} = P/T$ and $\frac{\partial S}{\partial N}|_{V,E} = -\mu/T$.

Whichever variables we hold constant in equation (3.75), the corresponding differentials are zero. This leads to the desired expressions. Note that the $|_{A,B,\dots}$ should NOT be omitted, as these keep track of which independent variables are being used.

(b) Show that, if f is a function of x and y , that $\frac{\partial x}{\partial y}|_f \frac{\partial y}{\partial f}|_x \frac{\partial f}{\partial x}|_y = -1$. Applying this formula to S at fixed E , derive the two equations in part (a) again.

Let's move along the path suggested in the problem: starting from (x_0, y_0, f_0) , moving at constant f to $y_0 + \Delta y$ brings us to the point:

$$\left(x_0 + \frac{\partial x}{\partial y}|_f \Delta y, y_0 + \Delta y, f_0 \right).$$

Moving then at constant x back to y_0 leads us to:

$$\left(x_0 + \frac{\partial x}{\partial y}|_f \Delta y, y_0, f_0 - \frac{\partial f}{\partial y}|_x \Delta y \right).$$

Finally at constant y back to (x_0, y_0) we come to:

$$\left(x_0, y_0, f_0 - \frac{\partial f}{\partial y}|_x \Delta y - \frac{\partial f}{\partial x}|_y \left(\frac{\partial x}{\partial y}|_f \Delta y \right) \right).$$

This must be identical with our starting point (assuming f is single-valued), whence:

$$\frac{\partial f}{\partial y}|_x = - \frac{\partial f}{\partial x}|_y \frac{\partial x}{\partial y}|_f.$$

Assuming that $f(x, y)$ is a smooth 1-1 function of y for each x , $\left. \frac{\partial f}{\partial y} \right|_x = \left(\left. \frac{\partial y}{\partial f} \right|_x \right)^{-1}$ and we obtain the desired result:

$$\left. \frac{\partial y}{\partial f} \right|_x \left. \frac{\partial f}{\partial x} \right|_y \left. \frac{\partial x}{\partial y} \right|_f = -1.$$

Applying this formula to $S = S(E, V)$ at fixed N and $S = S(E, N)$ at fixed V and making use of known thermodynamic relations for p , μ , and T yields the same equations as in part (a).

(c) Starting from the traditional definitions for P and μ , derive the two equations in part (a) again.

This is a straightforward application of our result in part (b)

3.11 Maxwell Relations.

(a) Use $E(S, V, N)$ to show the Maxwell relation $\frac{\partial T}{\partial V} S, N = -\frac{\partial P}{\partial S} V, N$. Generate two other similar formulas by taking other second partial derivatives of E . There are many of these relations.

$E(S, V, N)$ is a smooth function whose mixed second-order partial derivatives commute. For example, the Maxwell relation (3.69) follows from:

$$\frac{\partial^2 E}{\partial S \partial V} N = \frac{\partial^2 E}{\partial V \partial S} N \Rightarrow -\frac{\partial P}{\partial S} V, N = \frac{\partial T}{\partial V} S, N.$$

Others follow in exactly the same manner. Note that every one has conjugate functions on the diagonals (S and T , P and V), and that the numerator's conjugate is one of the two constants of each differentiation.

(b) Using equation 3.57, write formulas for $E(S, V, N)$, $T(S, V, N)$, and $P(S, V, N)$ for the ideal gas. Show explicitly that the Maxwell relation equation 3.70 is satisfied.

First invert equation (3.57) to find:

$$E(S, N, V) = \frac{3h^2 N^{\frac{5}{3}}}{4\pi m V^{\frac{2}{3}}} \exp \left[\frac{2S}{3Nk_B} - \frac{5}{3} \right].$$

Now use the formulas from part (a):

$$T = \frac{2E}{3k_B N} \text{ and } P = \frac{2E}{3V}.$$

We check the Maxwell relation by straightforward differentiation.

4.2 Liouville vs. the damped pendulum.

(a) In the steps leading from equation 4.5 to equation 4.7, why does Liouville's theorem not apply to the damped pendulum? More specifically, what is $\partial\dot{p}/\partial p$ and $\partial\dot{q}/\partial q$?

We have the Hamiltonian-like equations of motion,

$$\dot{q} = \frac{p}{M} \quad \dot{p} = -\gamma p - K \sin q.$$

In the proof of Liouville's theorem, we used $\partial\dot{p}/\partial p + \partial\dot{q}/\partial q = 0$, but clearly this is not the case, because

$$\frac{\partial\dot{p}}{\partial p} = -\gamma. \quad \frac{\partial\dot{q}}{\partial q} = 0$$

Since \dot{q} is independent of q , that term vanishes. How has this condition been broken? Recall that the condition that this sum vanishes originally came from the fact that the order of taking partial derivatives doesn't matter, so

$$\frac{\partial\dot{q}}{\partial q} \stackrel{?}{=} \frac{\partial^2 \mathcal{H}}{\partial q \partial p} = \frac{\partial^2 \mathcal{H}}{\partial p \partial q} \stackrel{?}{=} -\frac{\partial\dot{p}}{\partial p}.$$

But this assumes there exists \mathcal{H} for which $\dot{q} = \frac{\partial \mathcal{H}}{\partial p}$ and $\dot{p} = -\frac{\partial \mathcal{H}}{\partial q}$. In this case, such a function \mathcal{H} does not exist.

(b) Find an expression for the total derivative $d\rho/dt$ in terms of ρ for the damped pendulum. How does the probability density vary with time? If we evolve a region of phase space of initial volume $A = \Delta p \Delta q$ how will its volume depend upon time?

As an aside, we know that the long-time density will be

$$\rho_{t=\infty}(q, p) = \delta(p)\delta(q).$$

We begin with the definition of the total derivative,

$$\frac{d\rho}{dt} = \frac{\partial\rho}{\partial t} + \frac{\partial\rho}{\partial q}\dot{q} + \frac{\partial\rho}{\partial p}\dot{p}.$$

Now the continuity equation gives us

$$\frac{\partial\rho}{\partial t} = -\frac{\partial\rho\dot{q}}{\partial q} - \frac{\partial\rho\dot{p}}{\partial p} = -\frac{\partial\rho}{\partial q}\dot{q} - \frac{\partial\rho}{\partial p}\dot{p} - \rho\left(\frac{\partial\dot{q}}{\partial q} + \frac{\partial\dot{p}}{\partial p}\right).$$

Moving the first two terms from the right to the left allows us to write

$$\frac{d\rho}{dt} = -\rho\left(\frac{\partial\dot{q}}{\partial q} + \frac{\partial\dot{p}}{\partial p}\right) = \gamma\rho.$$

The solution is then exponential growth,

$$\rho(t) = e^{\gamma t}.$$

What does this mean? At any point the density at that point, as we follow the point's trajectory in phase space, is increasing. This makes sense, since neighboring trajectories are getting smooshed together with it. Since number is conserved, we see that an initial volume $A = \Delta p \Delta q$ will evolve as $A(t) = A \exp[-\gamma t]$.

4.3 Invariant Measures.

(a) **Iteration:** Set $\mu = 0.2$; iterate f for some initial points $0 < x_0 < 1$ of your choosing, and convince yourself that they are all attracted to zero. Plot f and the diagonal $y = x$ on the same plot. Are there any fixed points other than $x = 0$? Repeat for $\mu = 0.4$ and 0.6 . What happens?

Analytics: Find the non-zero fixed point $x^*(\mu)$ of the map 4.11, and show that it exists and is stable for $1/4 < \mu < 3/4$. If you're ambitious or have a computer algebra program, show that there is a stable period-two cycle for $3/4 < \mu < (1 + \sqrt{6})/4$.

See FIG. 22-26 for plots. We have a fixed point whenever $f(x) = x$. Therefore $x = 0$ is always a fixed point. Now we look at $f'(x) = 4\mu(1 - 2x)$. We see that when $\mu < 1/4$, $f'(0) < 1$ and $f''(x) < 0$ always. Therefore, $f(x) < x$ always and there are no other fixed points. When $\mu > 1/4$, there is a fixed point at $x_0 = 1 - 1/4\mu$. For stability,

$$|f(x) - x_0| < |x - x_0|.$$

We expand,

$$|f(x) - x_0| \approx |f(x_0) + f'(x_0)(x - x_0) - x_0| = |f'(x_0)(x - x_0)| < |x - x_0|$$

only if $|f'(x_0)| < 1$. We see that $f'(x_0) > -1$ for $\mu < 3/4$. Thus, there is a single nonzero stable fixed point for $1/4 < \mu < 3/4$.

For a two-cycle, we need $f \circ f$ to have a stable fixed point. We find that $(f \circ f)(x)$ has two additional fixed points:

$$x_{\pm} = \frac{1}{8\mu} \left(1 + 4\mu \pm \sqrt{16\mu^2 - 8\mu - 3} \right).$$

We can compute $(f \circ f)'(x_{\pm}) = -16\mu^2 + 8\mu + 4$ and find it to be astonishingly simple. We solve now $(f \circ f)'(x_{\pm}) > -1$ and see that we need $\mu < \frac{1+\sqrt{6}}{4}$.

(b) **Invariant Density:** Set $\mu = 1$; iterate f many times and form a histogram of values giving the density $\rho(x)$ of points along the trajectory. You should find that points x near the boundaries are approached more often than points near the center.

Analytics: Using the fact that the long time average $\rho(x)$ must be independent of time, verify for $\mu = 1$ that the density of points is $\rho(x) = 1/\pi\sqrt{x(1-x)}$. Plot this theoretical curve with your numerical histogram.

Numerics. See FIG. 27.

Analytics. First, let us provide an alternative, more formal derivation for eqn 4.15, giving the condition guaranteeing that the density is invariant. We know

$$\rho(x) = \int dx' \rho(x') \delta(f(x') - x).$$

Holding constant the x inside the δ -function, we rewrite it as

$$\delta(f(x') - x) = \sum_{x_i} \frac{\delta(x' - x_i)}{|f'(x_i)|},$$

where x_i are the roots of $f(x') = x$, namely

$$x_a, x_b = x_{\pm} = \frac{1}{2} \pm \frac{1}{2}\sqrt{1-x}.$$

Thus,

$$\rho(x) = \frac{\rho(x_+)}{|f'(x_+)|} + \frac{\rho(x_-)}{|f'(x_-)|},$$

giving eqn 4.15 with $x = f(x_a)$, $x_a = x_+$, and $x_b = x_-$.

Now, using eqn 4.15, we note that $f'(x_{\pm}) = \mp 4\sqrt{1-x}$ so the absolute values are the same. If we put in the given solution, $\rho(x) = 1/\pi\sqrt{x(1-x)}$ we find $\rho(x_{\pm}) = 2/\pi\sqrt{x}$. Combining this all, we see that indeed the condition is satisfied (and normalized) for this density.

(c) **Cusps in the invariant density:** Find the invariant density for $\mu = 0.9$. Make your trajectory length N_{cycles} big enough and the bin size small enough to see the interesting structures. Notice that the attractor no longer fills the whole range $(0, 1)$: locate roughly where the edges are. Notice also the cusps in $\rho(x)$ at the edges of the attractor, and also at places inside the attractor. Locate some of the more prominent cusps.

See FIG. 28.

(d) **Analytics of cusps:** Set $\mu = 0.9$. Calculate $f(\frac{1}{2})$, $f(f(\frac{1}{2}))$, \dots and compare these iterates to the locations of the edges and cusps from part (c).

See FIG. 28. The cusps line up with the images of $x = \frac{1}{2}$.

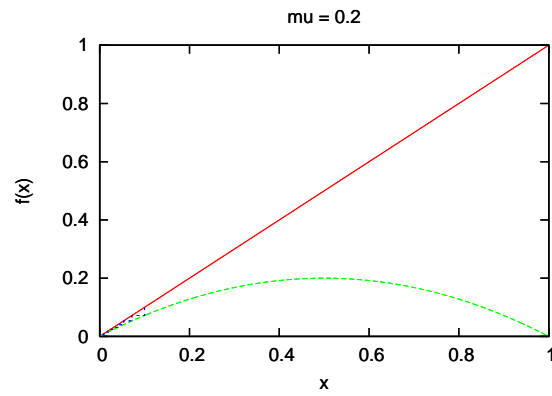
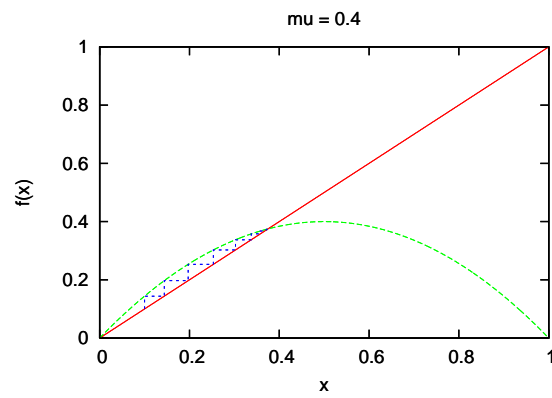
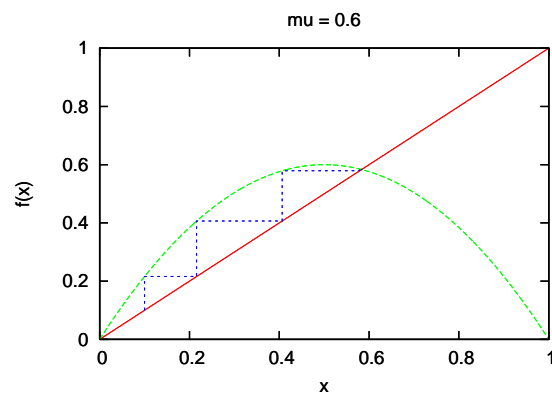
(e) **Bifurcation Diagram:** Plot the attractor as a function of μ , for $0.8 < \mu < 1$. On the same plot, for the same μ s, plot the first eight images of $x = \frac{1}{2}$, that is, $f(\frac{1}{2})$, $f(f(\frac{1}{2}))$, \dots . Are the boundaries you see just the cusps? What happens in the bifurcation diagram when two boundaries touch?

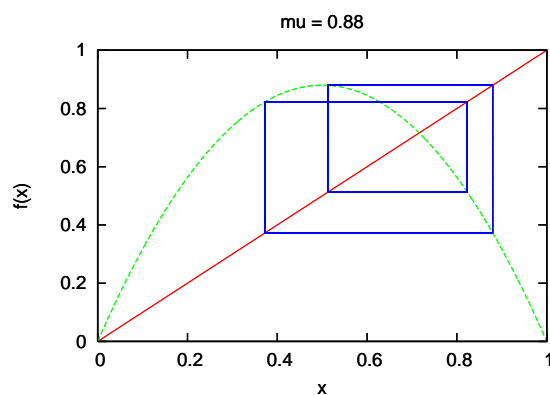
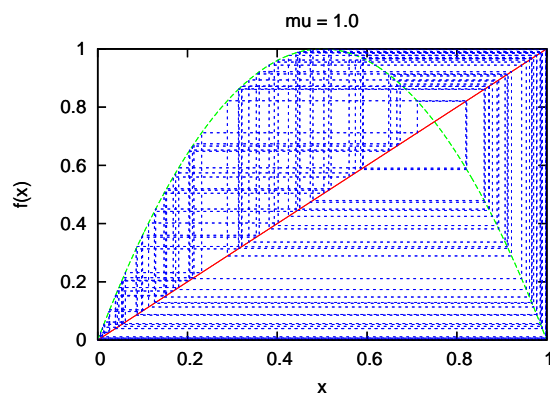
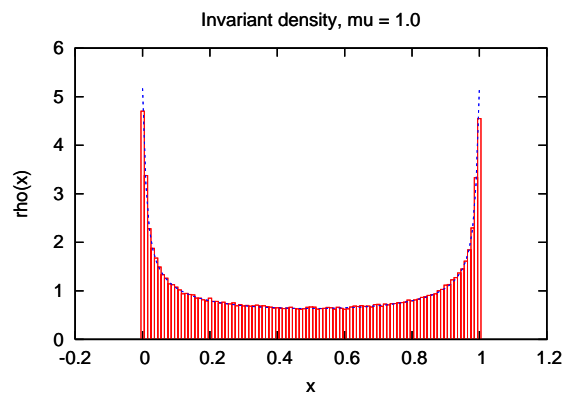
See FIG. 29. The boundaries on the bifurcation diagram, where the density jumps, are indeed the cusps in the invariant density for fixed μ . When two boundaries touch, one often sees a transition from chaos to periodic behavior. (This makes sense. The boundaries are images of the point $x = \frac{1}{2}$, so the intersection of two boundaries means that $f^n(\frac{1}{2}) = f^m(\frac{1}{2})$ for some $n > m > 0$, meaning that $f^{n-m}(y) = y$ for $y = f^m(\frac{1}{2})$, and hence one has a periodic orbit.)

4.4 Jupiter! and the KAM Theorem.

(a) Estimate the fraction that the Earth's radius from the Sun changes during the first Jovian year. How much does this fractional variation increase over the next hundred Jovian years?

These should be small numbers on the order of a few percent. The fractional variation does not increase significantly with time.

FIG. 22: Map and identity for $\mu = 0.2$.FIG. 23: Map and identity for $\mu = 0.4$.FIG. 24: Map and identity for $\mu = 0.6$. Note the stable fixed point.

FIG. 25: Map and identity for $\mu = 0.88$. Note the period-four cycle.FIG. 26: Map and identity for $\mu = 1.0$. Chaos!FIG. 27: Invariant density for $\mu = 1.0$, fit to theoretical curve.

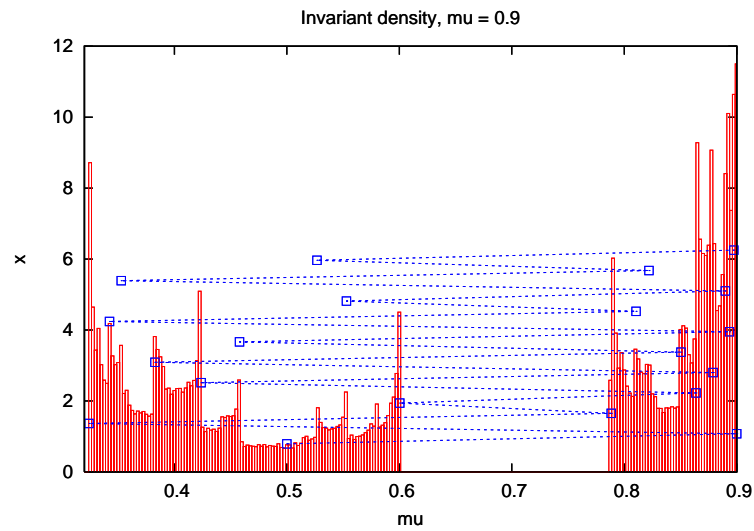
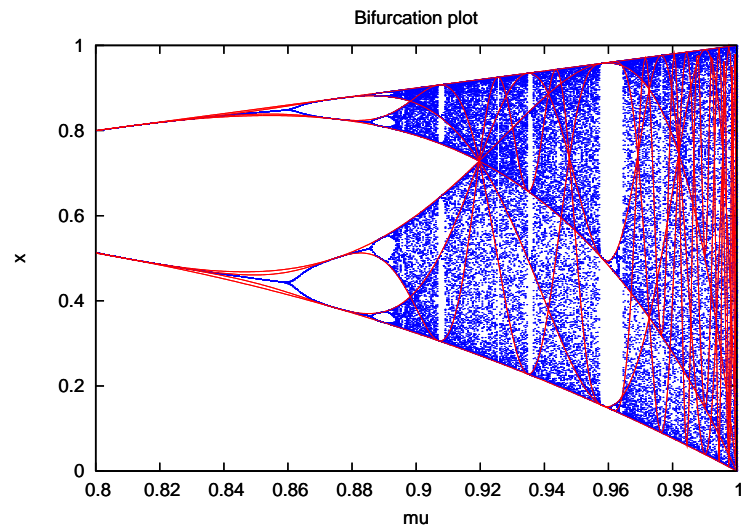
FIG. 28: Invariant density and first 20 iterates for $\mu = 0.9$.

FIG. 29: Bifurcation plot and first eight iterates.

(b) In the non-interacting planet approximation, what topological surface is it in in the eighteen-dimensional phase space that contains the trajectory of the three bodies? About how many times does Earth wind around this surface during each Jovian year?

Non-interacting topological surface: a torus. The Earth's trajectory in its six-dimensional phase space is a closed loop, which is topologically equivalent to a circle. The same goes for Jupiter. The Sun sits at the origin in its phase space and doesn't contribute to the topology. $S(1) \otimes S(1) = \text{a torus}$. The Earth orbits the Sun about 11.9 times during each Jovian year.

(c) About how large can you make Jupiter's mass before Earth's orbit stops looking like a torus? Admire the cool orbits when the mass becomes too heavy.

Many answers for this one. Say, 20000 Earth masses?

(d) If you can do a screen capture, print out a Poincaré section with initial conditions both on KAM tori and in chaotic regions: label each.

Pretty pictures...

5.1 Life and the Heat Death of the Universe.

(a) What is the minimum energy Q needed per thought, in terms of ΔS and T_2 ? You may take T_1 to be very large.

For the case of minimum energy per thought, no work (whatever that would mean) is done in the process. Therefore $Q = Q_1 = Q_2$ and $\Delta S = \Delta S_1 + \Delta S_2 = Q(1/T_2 - 1/T_1)$. If we take T_1 to be very large, then $\Delta S \approx Q/T_2$ and $\boxed{Q_{min} \approx T_2 \Delta S}$ is the minimum energy needed per thought.

(b) Write an expression for the maximum rate of thoughts per unit time dH/dt (the inverse of the time Δt per thought), in terms of ΔS , C , and T_2 .

For the maximum rate of thoughts dH/dt , the energy consumed per thought needs to be minimum while the power radiated is maximum: $P_{max} = CT_2^3$. Then

$$\left. \frac{dH}{dt} \right|_{max} = \frac{1}{\Delta t_{min}} = \frac{P_{max}}{Q_{min}} = \frac{cT_2^3}{\Delta S T_2} = \frac{cT_2^2}{\Delta S}.$$

(c) How many thoughts H can an ecologically efficient being have between now and time infinity, in terms of ΔS , C , A , and the current time t_0 ?

With $T_2 = \Theta(t) = A/t$, the number of thoughts that an ‘ecologically efficient’ being could have is

$$H_{max} = \int_{t_0}^{\infty} dt \frac{dH}{dt} \Big|_{max} = \int_{t_0}^{\infty} dt \frac{c T_2^3}{\Delta S} = \frac{c A^2}{\Delta S t_0},$$

which is finite.

(d) Show that with Dyson’s cooling schedule, the total number of thoughts H is infinite, but the total energy consumed U is finite.

If now $T_2(t) = A t^{-3/8}$, the maximum number of thoughts is

$$H_{max} = \int_{t_0}^{\infty} dt \left(\frac{c}{\Delta S} A^2 t^{-3/4} \right) \rightarrow \infty,$$

while the corresponding energy consumed would be

$$U = \int_{t_0}^{\infty} dt \frac{dQ_1}{dt} = \int_{t_0}^{\infty} dt Q_{min} \frac{dH}{dt} \Big|_{max} = \dots = 8c A^3 t_0^{-1/8},$$

which is finite.

5.2 Burning Information and Maxwellian Demons .

(a) How much work is done by the atom as the piston retracts?

$$W = k_B T \ln(2)$$

(b) Explain how to convert one bit into another.

Lower the upper piston; Remove the partition; Move both pistons upwards slowly; introduce partition; lower lower piston.

(c) What is the change in entropy if the volume doubles?

$$\Delta S = k_B \ln(2), \text{ and } \Delta S_s = 1.$$

(d) Demonic States

The demon still remembers the (now erased) bit of information on the tape. This information is now garbage—unrelated to anything in the outside world. In this sense the demon is in an unknown state, with entropy $k_B \log 2$. It would take work $k_B T \log(2)$ to return the demon’s memory to a known, useful state for doing another measurement, just as in part (a). There is hence no violation of the second law.

5.3 Reversible Computation.

(a) Irreversible logic gates.

One bit is lost during the operation of the irreversible logic gate (initially, $S_I = 2$, afterwards $S_I = 1$. The change in thermodynamic entropy is $-k_B \ln(2)$, so the work necessary to dump the excess entropy in the environment is $k_B T \ln(2)$.

(b) Reversible Gate.

A, B	$C, D = A', B'$	C', D'
11	10	11
10	11	10
01	01	01
00	00	00

Hence, two controlled-not gates equal the identity operation.

5.4 Black Hole Thermodynamics.

(a) Calculate the specific heat of the black hole.

We rewrite the temperature and energy in terms of mass,

$$T_{bh} = \frac{\hbar c^3}{8\pi M G k_B} \quad E = M c^2.$$

Now

$$\frac{1}{C} = \frac{\partial T}{\partial E} = \frac{\partial T}{c^2 \partial M} = -\frac{\hbar c}{8\pi G k_B} \frac{1}{M^2}.$$

The specific heat is thus

$$C = -\frac{8\pi k_B G}{\hbar c} M^2 = -\frac{\hbar c^5}{8\pi G k_B} \frac{1}{T^2}.$$

(b) Calculate the entropy of the black hole, by using the definition of temperature, $1/T = \frac{\partial S}{\partial E}$ and assuming the entropy is zero at mass $M = 0$. Express your results in terms of the surface area $A = 4\pi R_s^2$, measured in units of the Planck length, $L^* = \sqrt{\hbar G/c^3}$.

We integrate $1/T = \partial S/\partial E$ to find

$$S(M) - S(0) = \int_0^{Mc^2} \frac{1}{T} dE = \int_0^{Mc^2} \frac{8\pi G k_B}{\hbar c^5} (Mc^2) dE = \frac{4\pi k_B G (Mc^2)^2}{\hbar c^5}.$$

Since we define $S(M) = 0$ at $M = 0$, we find

$$S = \frac{4\pi k_B G M^2}{\hbar c}.$$

We finally rewrite this in natural units,

$$S = \frac{4\pi k_B G}{\hbar c} \left(\frac{R_s^2 c^4}{4G^2} \right) = k_B \frac{A}{4},$$

where $A = 4\pi (R/L^*)^2$.

(c) Calculate the maximum number of bits that can be stored in a sphere of radius one centimeter.

The entropy of a bit is $S_{\text{bit}} = k_B \log 2$. The most information we can have in a 1 cm^2 sphere is

$$S_{\text{max}} = k_B \pi \left(\frac{c^3}{\hbar G} \right) (1\text{cm})^2.$$

This gives

$$\frac{S_{\text{max}}}{S_{\text{bit}}} = 1.74 \times 10^{66} \text{ bits.}$$

5.5 P-V Diagram.

Which of the following are true?

(1) The cycle is reversible: no net entropy is created in the universe.

False. Leg a has a warm system exchanging heat irreversibly with a cold bath, and leg b has it touching a hot bath. Hence a net entropy is generated during a cycle.

(2) The cycle acts as a refrigerator, using work from the piston to draw energy from the cold bath into the hot bath, cooling the cold bath.

False. The piston does not cool the cold bath. Only in leg a is there a heat exchange with the cold bath, and heat is entering the cold bath during that leg. (The net energy $E = (3/2)PV$ goes down in the piston during that leg.)

(3) The cycle acts as an engine, transferring heat from the hot bath to the cold bath and doing positive net work on the outside world.

False. The net work done by the piston is the area $\int P dV < 0$ swept out by the curve. It is not an engine, because it absorbs work.

(4) The work done per cycle has magnitude $|W| = P_0 V_0 |4 \log 4 - 3|$.

True. The net work done by the outside world on the piston is $\int P_c - P_b dV = -\int (4P_0V_0/V - P_0) dV = -P_0V_0(4[\log 4V_0 - \log V_0]) - 3) = -P_0V_0(4\log 4 - 3)$.

(5) The heat transferred into the cold bath, Q_c has magnitude $|Q_c| = (9/2)P_0V_0$.

True. In leg a (the only one which matters), the gas in the piston has energy $U = (3/2)Nk_B T = (3/2)PV$. It thus starts with energy $(3/2)(4P_0V_0) = 6P_0V_0$ and ends with energy $(3/2)P_0V_0$, so a net energy $(9/2)P_0V_0$ enters the cold bath.

(6) The heat transferred from the hot bath Q_h , plus the net work W done by the piston onto the gas, equals the heat Q_c transferred into the cold bath.

True. This follows from energy conservation and the fact that the system returns to its initial state after a cycle.

5.6 Carnot Refrigerator.

(a) How many watts of energy leak from our refrigerator through this insulation?

This is straightforward. Adding up the faces, $A = 10\text{m}^2$, so that

$$\dot{Q} = \frac{\kappa A \Delta T}{d} = \frac{(0.02\text{W/mK})(10\text{m}^2)(30\text{K})}{0.03\text{m}} = 200\text{W}.$$

(b) Suppose (i) we don't open the refrigerator door, (ii) the thermal losses are dominated by the leakage through the foam and not through the seals around the doors, and (iii) the refrigerator runs a perfectly efficient Carnot cycle. How much power on average will our refrigerator need to operate? What fraction of the time will the motor run?

We are told that $V = 120\text{V}$ and $I = 4.75\text{A}$ so that $P = 570\text{W}$. We know for a Carnot cycle that $\Delta S = 0$ so that $\frac{|Q_h|}{T_h} = \frac{|Q_c|}{T_c}$. This also holds for power. The refrigerator has 200W leaking in, so the motor must remove $\dot{Q}_c = 200\text{W}$ from the refrigerator (cold bath) to maintain its temperature. Thus, $\dot{Q}_h = -\frac{300\text{K}}{270\text{K}}200\text{W} = -222\text{W}$. And we have $\dot{W} = \dot{Q}_h + \dot{Q}_c = -22\text{W}$. Let t be total time elapsed, and t' be the time the motor runs. Then $W = (570\text{W})t' = (22\text{W})t$ so that

$$t' = \frac{22}{570}t.$$

The motor runs $22/570 \approx 3.9\%$ of the time.

5.7 Does Entropy Increase?.

Show for any function $f(\rho)$ that $\partial f(\rho)/\partial t = -\nabla \cdot [f(\rho)\mathbb{V}] = -\sum_{\alpha} \partial/\partial p_{\alpha}(f(\rho)\dot{p}_{\alpha}) + \partial/\partial q_{\alpha}(f(\rho)\dot{q}_{\alpha})$, where $\mathbb{V} = (\dot{\mathbb{P}}, \dot{\mathbb{Q}})$ is the $6N$ dimensional velocity in phase space. Hence, show $\int \partial f(\rho)/\partial t d\mathbb{P}d\mathbb{Q} = 0$, assuming that the probability density vanishes at large momenta and positions and $f(0) = 0$. Show, thus, that the entropy $S = -k_B \int \rho \log \rho$ is constant in time.

Let $\rho = \rho(q_{\alpha}, p_{\alpha}, t)$ be a probability distribution in phase space and $\rho(t) = \rho(q_{\alpha}(t), p_{\alpha}(t), t)$ be the distribution along a particular trajectory or streamline given by $q_{\alpha}(t)$ and $p_{\alpha}(t)$. Then for any $f(\rho)$ satisfying the assumptions described in the problem:

$$\frac{\partial f(\rho)}{\partial t} \equiv \frac{\partial}{\partial t} f(\rho(q_{\alpha}, p_{\alpha}, t)) = \frac{df(\rho)}{d\rho} \frac{\partial}{\partial t} \rho(q_{\alpha}, p_{\alpha}, t).$$

Liouville's theorem applies along the streamline:

$$\frac{d\rho(t)}{dt} \equiv \frac{d}{dt} \rho(q_{\alpha}(t), p_{\alpha}(t), t) = \frac{\partial}{\partial t} \rho(q_{\alpha}, p_{\alpha}, t) + (\nabla \rho(q_{\alpha}, p_{\alpha}, t)) \cdot \mathbb{V} = 0.$$

We therefore have:

$$\frac{\partial f(\rho)}{\partial t} = - \left(\frac{df(\rho)}{d\rho} \nabla \rho \right) \cdot \mathbb{V} = -(\nabla f) \cdot \mathbb{V} = -\nabla \cdot [f\mathbb{V}],$$

where the last step holds because $\nabla \cdot \mathbb{V} = \sum_{\alpha} \partial \dot{p}_{\alpha} / \partial p_{\alpha} + \partial \dot{q}_{\alpha} / \partial q_{\alpha} = 0$ for a Hamiltonian system. (This follows from Hamilton's equations of motion, eqn 4.6). Using Gauss's law:

$$\int_{\text{phase space}} \frac{\partial f(\rho)}{\partial t} d\mathbb{P} d\mathbb{Q} = \oint_{\text{surface at } R \rightarrow \infty} f(\rho) \mathbb{V} \cdot d\vec{a} = 0,$$

assuming $\rho(\mathbb{P}, \mathbb{Q}) = 0$ for large \mathbb{P} and/or \mathbb{Q} (i.e. ρ has compact support). Now let's look at the rate of change of the entropy:

$$\frac{dS}{dt} = -k_B \int \frac{\partial}{\partial t} (\rho \log \rho) d\mathbb{Q} d\mathbb{P}.$$

Taking $f(\rho) = \rho \log \rho$, we see that $f(\rho) \rightarrow 0$ as $\rho \rightarrow 0$. The above result therefore applies:

$$\frac{dS}{dt} = 0.$$

The entropy is a constant in time.

5.8 The Arnol'd Cat.

Check that Γ preserves area. Show that it takes a square $n \times n$ and maps it onto itself with periodic boundary conditions. As a linear map, find the eigenvalues and eigenvectors. Argue that a small neighborhood will initially be stretched along an irrational direction into a thin strip.

Ignoring the mod n part, Γ is a non-singular linear map that can be represented as multiplication by the matrix $M = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$.

The determinant of the matrix is equal to the Jacobian of the transformation—the factor that multiplies the transformed element of area (when changing variables in an integral, e.g.). Since $\det M = 1$, area is preserved in this case. We should be careful though, because with the mod n back in the picture, we need to check that the mapping is still one-to-one in

order to guarantee that it preserves area. Let (x, y) and (u, v) be two pairs of points in the $n \times n$ square $x, y \in [0, n)$. Then

$$\Gamma \begin{pmatrix} x \\ y \end{pmatrix} = \Gamma \begin{pmatrix} u \\ v \end{pmatrix} \iff \begin{cases} 2x + y \equiv 2u + v \pmod{n} \\ \text{and } x + y \equiv u + v \pmod{n} \end{cases}.$$

Noting that $x, y, u, v \in [0, n)$, it follows that $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix}$ and Γ is area preserving. Since Γ is one-to-one and area preserving, it maps the $n \times n$ square onto itself with periodic boundary conditions (due to the mod n). As a linear map (just the matrix M), we solve for the eigenvalues and eigenvectors in the usual way to find:

$$\text{eigenvalues } \lambda_{\pm} = \frac{3 \pm \sqrt{5}}{2} \text{ and eigenvectors } \vec{v}_{\pm} = \begin{pmatrix} (1 \pm \sqrt{5})/2 \\ 1 \end{pmatrix}.$$

The mapping can be rearranged to look like UVU^{-1} with U the matrix of (normalized) eigenvectors and V the diagonal eigenvalues. The action of the map in this form is seen to be an expansion by a factor of $\lambda_+ > 1$ along the first eigenvector and a contraction by a factor of $\lambda_- < 1$ along the second. A small region around $(x, y) = (0, 0)$ will therefore initially be stretched in an irrational direction since the eigenvector \vec{v}_+ has irrational slope.

5.9 Chaos, Lyapunov, and Entropy Increase.

Start with $\mu = 0.9$ and two nearby points x_0 and $y_0 = x_0 + \epsilon$ somewhere between zero and one. Investigate the two trajectories $x_0, f(x_0), f(f(x_0)), \dots, f^{[n]}(x_0)$ and $y_0, f(y_0), \dots$. How fast do they separate? Estimate the Lyapunov exponent.

See FIG. 30,31. We find $\lambda \approx 0.2$ from the slope of the fit line, so that $|x - y| \sim \epsilon e^{\lambda t}$

5.10 Entropy Increases: Diffusion.

Diffusion Equation Entropy. Derive a formula for the time derivative of the entropy $S = -k_B \int \rho(x) \log \rho(x) dx$ and show that it strictly increases in time.

By the product rule, the time derivative of the entropy is

$$\frac{dS}{dt} = -k_B \int \left(\frac{\partial \rho}{\partial t} \log \rho + \frac{\partial \rho}{\partial t} \right) dx.$$

We can pull the time-derivative out of the second term of the integral, and substitute the diffusion equation $\partial \rho / \partial t = D \partial^2 \rho / \partial x^2$ in the first term:

$$\frac{dS}{dt} = -k_B \int D \frac{\partial^2 \rho}{\partial x^2} \log \rho dx - \frac{d}{dt} \int \rho dx.$$

Now the second term is simply $\frac{d}{dt} 1$, since ρ is conserved. The first term is integrated by parts, where the boundary term, $\partial \rho / \partial x \log \rho$, vanishes at $x = \pm \infty$.

$$\frac{dS}{dt} = k_B D \int \left(\frac{\partial \rho}{\partial x} \right)^2 \frac{dx}{\rho} > 0$$

since the integrand is positive definite.

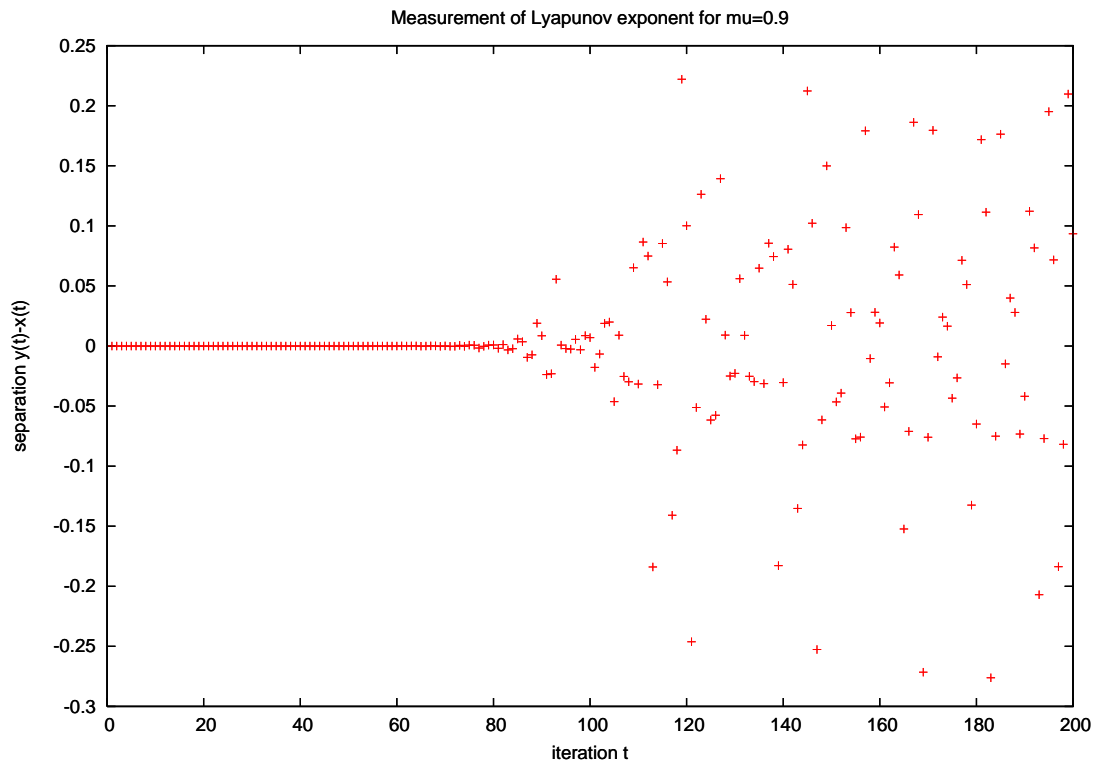


FIG. 30: Separation, linear scale

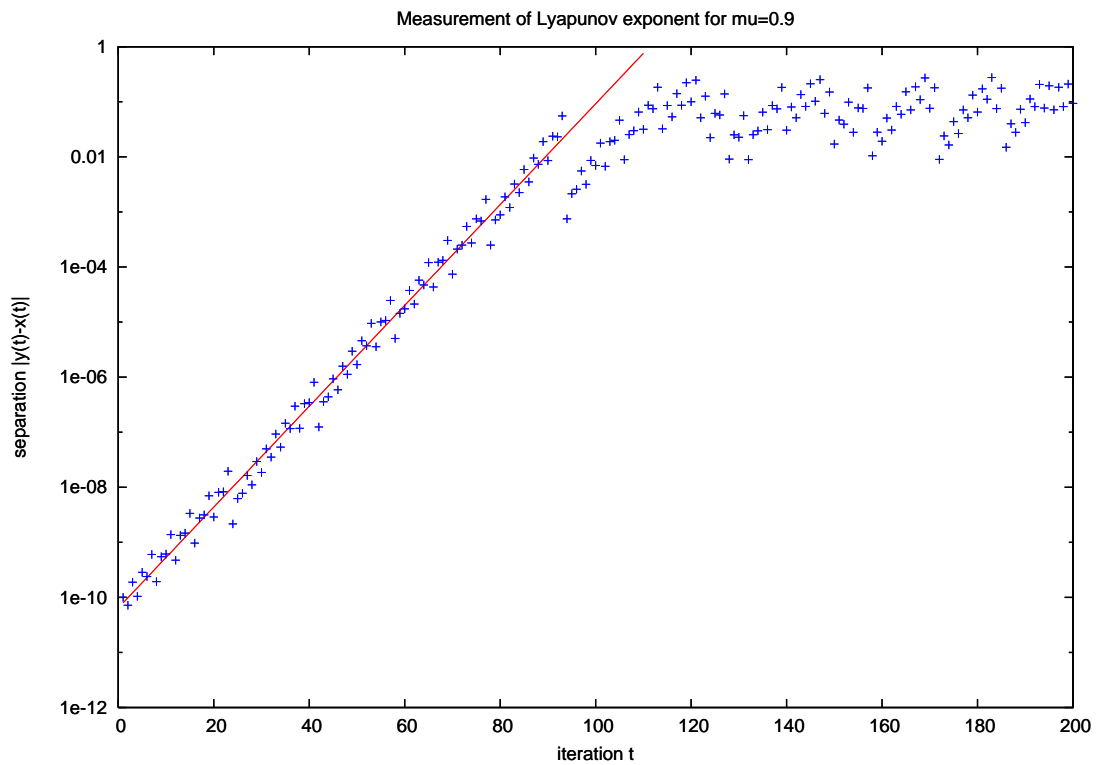


FIG. 31: Separation, semilog scale. Notice the fit line which shows an exponential sensitivity until the system reaches chaos.

5.11 Entropy of Glasses.

(a) Is the residual entropy, equation 5.42, larger on heating or on cooling?

Referring to Fig. 5.13 in the text, the entropy flow

$$\int_0^{T_1} \frac{1}{T} \frac{dQ}{dT} dT = \int_0^{T_1} C_p/T dT$$

is smaller for the heating curve since weight is shifted to higher T (hence smaller C_p/T). Therefore, the residual entropy on heating (S_{heat}) is greater than on cooling (S_{cool}).

(b) By using the second law, show that when cooling and then heating from an equilibrium liquid the residual entropy measured on cooling must always be less than the residual entropy measured on heating.

Consider the entropy flow from the glass into the outside world (bath) on cooling (ΔS_{cool}) and the flow into the glass on subsequent re-heating (ΔS_{heat}). Consider a cycle from temperature T_2 (liquid glass) to T_1 ('solid' glass) and back to T_2 . The entropy of the equilibrium liquid glass is well-defined, and our system begins and ends in the same state. The net entropy change of the glass is therefore zero and the total entropy change for the cycle is

$$\Delta S_{\text{total}} = \Delta S_{\text{bath}} = \Delta S_{\text{cool}} - \Delta S_{\text{heat}} \geq 0,$$

by the second law of thermodynamics. Therefore, $\Delta S_{\text{heat}} \leq \Delta S_{\text{cool}}$, and just as in part (a), this implies that $S_{\text{cool}} \leq S_{\text{heat}}$.

(c) In a molecular dynamics simulation with one hundred indistinguishable atoms, and assuming that the residual entropy is $k_B \log 2$ per atom, what is the probability that two coolings to zero energy will arrive at equivalent atomic configurations (up to permutations)? In a system with 10^{23} molecular units, with residual entropy $k_B \log 2$ per unit, about how many coolings would be needed to arrive at the original configuration again, with probability $1/2$?

A residual entropy of $k_B \log 2$ per atom indicates that there are 2^N nonequivalent states the glass can end up in. The probability of landing in any particular one is therefore $p = 1/2^N$. The probability of arriving at any configuration twice in two coolings is the same (the first cooling can be anything, while the second must be equivalent). Because we are given the entropy, we needn't worry about complications and $N!$ terms from the identical particles.

For $N = 10^{23}$ atoms, the probability of arriving at "the original configuration" again would be $p = 1/2^N$, with $N = 10^{23}$. The number of coolings needed to arrive at the original configuration is thus roughly also $2^N = 2^{10^{23}} \sim 10^{3 \times 10^{22}}$.

5.12 Rubber Band.

(a) Find an exact formula for the entropy of this system in terms of d , N , and L .

We simply count the states. Suppose we have n_+ rightward links and n_- leftward, so that $n_+ + n_- = N$ is the total number of links and $(n_+ - n_-)d = L$ is the total length, so $n_+ = N/2 + L/2d$. The number of configurations in this constant-length ensemble is then

$$\Omega = \binom{N}{n_+} = \frac{N!}{n_+! n_-!} = \frac{N!}{(N/2 + L/2d)! (N/2 - L/2d)!}.$$

Taking the log,

$$S = k_B \log \Omega = k_B [\log N! - \log (N/2 + L/2d)! - \log (N/2 - L/2d)!].$$

(b) Find an expression for the force $-F$ exerted by the bath on the molecule in terms of the bath entropy. Using the fact that the length L must maximize the entropy of the universe, write a general expression for F in terms of the internal entropy S of the molecule.

We have the relationship $dE_{\text{bath}} = (-F)dL$. Thus,

$$\frac{1}{T} = \frac{\partial S_{\text{bath}}}{\partial E_{\text{bath}}} = -\frac{1}{F} \frac{\partial S_{\text{bath}}}{\partial L}.$$

But we know that L maximizes S_{univ} so that

$$\frac{\partial S_{\text{univ}}}{\partial L} = \frac{\partial S_{\text{bath}}}{\partial L} + \frac{\partial S_{\text{band}}}{\partial L} = 0.$$

Therefore,

$$F = T \frac{\partial S_{\text{band}}}{\partial L}.$$

(c) Take our model of the molecule from part (a), the general law of part (b), and Stirling's formula, write the force law $F(L)$ for our molecule for large lengths N . What is the spring constant K in Hooke's law $F = -KL$ for our molecule, for small L ?

We recall the result of (a) and apply Stirling's formula,

$$\begin{aligned} S &\approx k_B [N \log N - (N/2 + L/2d) \log (N/2 + L/2d) - (N/2 - L/2d) \log (N/2 - L/2d)] \\ &= \frac{k_B}{2d} \left[Nd \log \frac{4N^2 d^2}{(Nd + L)(Nd - L)} - L \log \frac{Nd + L}{Nd - L} \right]. \end{aligned}$$

Now we differentiate,

$$\begin{aligned} F &= T \frac{k_B}{2d} \left[Nd \frac{\partial}{\partial L} \log \frac{4N^2 d^2}{(Nd + L)(Nd - L)} - L \frac{\partial}{\partial L} \log \frac{Nd + L}{Nd - L} - \log \frac{Nd + L}{Nd - L} \right] \\ &= -\frac{k_B T}{2d} \log \frac{Nd + L}{Nd - L}. \end{aligned}$$

Finally we Taylor expand $L \ll Nd$,

$$F(L) \approx -\frac{k_B T}{2d} \log \left(1 + 2 \frac{L}{Nd} \right) \approx -\frac{k_B T}{Nd^2} L = -KL,$$

where $K = \frac{k_B T}{Nd^2}$.

(d) If we increase the temperature of our rubber band while it is under tension, will it expand or contract?

If we increase the temperature, the rubber band will contract. The lower entropy stretched-out states become “more expensive.” We can try to understand this as follows, recalling the partial derivative identity we learned last week,

$$\left. \frac{\partial L}{\partial T} \right|_F = - \left. \frac{\partial L}{\partial F} \right|_T \left. \frac{\partial F}{\partial T} \right|_L.$$

Now we have a Maxwell relation

$$\left. \frac{\partial F}{\partial T} \right|_L = \left. \frac{\partial S}{\partial L} \right|_T = \frac{F}{T} = - \frac{KL}{T}$$

and the spring constant,

$$\left. \frac{\partial L}{\partial F} \right|_T = - \frac{1}{K}.$$

We combine these so that

$$\left. \frac{\partial L}{\partial T} \right|_F = - \frac{L}{T} < 0$$

Thus the rubber band contracts with increasing temperature and constant force.

(e) True or false? ...

FFFTTT. The first four are more or less exclusive (not quite). When we stretch the rubber band, the configurational entropy decreases so the vibrational entropy increases, increasing the temperature. Inversely, relaxing the band will increase configurational entropy, decreasing vibrational entropy and temperature.

We have seen that work on the system (changing the length) causes a temperature change. But we notice that it is backwards from an ideal gas, where compression cools and expansion heats. This is because the rubber's force is inwards, while pressure is outward. That is, $dW_{\text{gas}} = PdV$ while $dW_{\text{band}} = -FdL$.

5.13 How Many Shuffles?.

(a) What is the information entropy of the deck before it is shuffled? After it is completely randomized?

Before being shuffled, we know exactly what the order of the deck is, so its entropy is zero. After being completely randomized:

$$S_s = \frac{\ln(52!)}{\ln(2)} \simeq 226$$

(b) What is the information entropy after one riffle, ignoring redundancies?

The riffle shuffle makes each of the 2^{52} states equally likely, hence:

$$S_s = \frac{\ln(2^{52})}{\ln(2)} = 52$$

(c) How many of the riffles drop the entire bottom half and then the entire top half, leaving the card ordering unchanged? What is the actual information entropy?

There are 52 ways of cutting the deck. Once the deck is cut, there is only one way to reproduce the original ordering, thus 52 riffles yield the same ordering. Thus, the actual entropy after one riffle is:

$$S_s = \frac{\ln(2^{52} - 52)}{\ln(2)} = 52 + \Delta S$$

$$\Delta S \simeq -1.7 \times 10^{-14}$$

the change is negligible!

(d) How many shuffles does it take to reach the entropy of a perfectly randomized deck, neglecting redundancies?

The “neglecting redundancies” assumption gives us a constant information entropy gain of 52 for every shuffle. Thus, we need n shuffles such that $52 \times n > 226$.

In this approximation, it takes 5 riffle shuffles. (Neglecting redundancies clearly becomes a serious approximation at this point.)

5.15 Shannon entropy.

(a) What is the Shannon entropy in this language? More specifically, what is the Shannon entropy rate?

The Shannon entropy of the ensemble of one letter messages (i.e. entropy per sound/letter) is

$$S = -k_B \sum_i^N p_m \log p_m = -\frac{1}{\log 2} \left\{ \frac{1}{2} \log \frac{1}{2} + \frac{1}{4} \log \frac{1}{4} + \frac{1}{4} \log \frac{1}{4} \right\} = \frac{3}{2}.$$

(b) Show that a communication channel transmitting bits can transmit no more than one unit of Shannon entropy per bit.

First we prove that for $N = 2^m$ messages, the Shannon entropy is extremized by $p_m = 1/N$. If we wish to find the extrema of S subject to the constraint $\phi(p_m) = \sum_m p_m - 1 = 0$, one way to do this is by the method of undetermined Lagrange multipliers. We introduce a new variable λ for each constraint (only one in this case) and solve the unconstrained problem of extremizing $g(p_m, \lambda) \equiv S(p_m) + \lambda \phi(p_m)$.

$$\frac{\partial g}{\partial \lambda} = 0 \text{ reproduces the constraint equation.}$$

$$\frac{\partial g}{\partial p_i} = \frac{\partial S}{\partial p_i} + \lambda \frac{\partial \phi}{\partial p_i} = -k_B (\log p_i + 1) + \lambda = 0, \quad i = 1, \dots, N.$$

Therefore,

$$p_i = \exp(\lambda/k_B - 1) = \text{constant} = \frac{1}{N},$$

where the last step is by the constraint condition (normalization). We see that the extremum we found is in fact a maximum, because there is only one extremum (as we found) and we can guess a distribution for which the entropy is less. In particular, we note that with $\rho = 1/N = 1/2^m$, $S = m$ is greater than zero, while picking one $p_k = 1$ and the rest 0 would give $S = 0$. For the case of one bit, there are only two possible messages: 0 or 1. The maximum entropy is obtained when the probabilities are equal: $p_0 = p_1 = 1/2$, yielding an entropy (per bit) of 1.

(c) In general, argue that the Shannon entropy gives the minimum number of bits needed to transmit the ensemble of messages.

The Shannon entropy depends on the number of messages in an ensemble and their respective probabilities, not on how or with how many bits the messages are encoded. So for a given ensemble of messages, if we encode it using the minimum number of bits, we will have the maximum entropy per bit. Since the maximum entropy per bit is one, the minimum number of bits necessary to encode the ensemble is precisely equal to the Shannon entropy. For the case of A'bc! communication, the minimum number of bits required per letter on average is $3/2$.

(d) Find a message compression scheme that is optimal, in the sense that it saturates the bound you derived in part (b).

One possible compression scheme is:

$$C \rightarrow 0, \quad A \rightarrow 10, \quad \text{and} \quad B \rightarrow 11.$$

With this scheme, the average number of bits per letter is

$$1p_C + 2p_A + 2p_B = \frac{3}{2},$$

and the probability of a zero or one is $p_0 = p_1 = 1/2$.

5.17 Deriving Entropy.

(a) For any rational probabilities q_ℓ , let g be the least common multiple of their denominators, and let $q_\ell = g_\ell/g$ for integers g_ℓ . Show that

$$S_I(B) = L(g) - \sum_{\ell} q_{\ell} L(g_{\ell}).$$

We see that $S_I(AB) = S_I(1/g, 1/g, \dots) = L(g)$. We use the third condition, writing it backwards to solve for $S_I(B)$:

$$\begin{aligned} S_I(B) &= S_I(AB) - \langle S_I(A|B_{\ell}) \rangle_B \\ &= L(g) - \sum_{\ell} q_{\ell} S_I(c_{1\ell}, \dots, c_{\Omega\ell}) \end{aligned}$$

We see from the figure that $c_{k\ell} = 1/g_\ell$ so that $S_I(c_{1\ell}, \dots, c_{\Omega\ell}) = L(g_\ell)$ and therefore

$$S_I(B) = L(g) - \sum_{\ell} q_{\ell} L(g_{\ell}).$$

(b) If $L(g) = k_S \log g$, show that equation 5.54 is the Shannon entropy 5.23.

We substitute $q_{\ell} = g_{\ell}/g$ so that

$$\begin{aligned} S_I(B) &= -k_S \left(-\log g + \sum_{\ell} q_{\ell} \log(g_{\ell}) \right) \\ &= -k_S \left(-\sum_{\ell} q_{\ell} g + \sum_{\ell} q_{\ell} \log(g_{\ell}) \right) \\ &= -k_S \sum_{\ell} q_{\ell} \log q_{\ell} \end{aligned}$$

where we inserted the sum $\sum_{\ell} q_{\ell} = 1$ since g doesn't depend on ℓ , and then combined the logs with $q_{\ell} = g_{\ell}/g$. This is Shannon entropy.

(c) Show that $L(g)$ is monotone increasing with g .

Abandoning the result from (b), we begin by comparing $L(a)$ and $L(b)$ with integers $a < b$. We have by definition

$$L(a) = S_I(1/a, \dots, 1/a) \quad L(b) = S_I(1/b, \dots, 1/b).$$

We use the second property to extend $L(a)$ to have the same number of arguments,

$$L(a) = S_I(\underbrace{1/a, \dots, 1/a}_a, \underbrace{0, \dots, 0}_{b-a}).$$

So by the first principle, $L(b) > L(a)$ since they have the same number of arguments, but the distribution in b is uniform. Thus, as x increases, $L(x)$ increases, so L is monotonic.

(d) Show $L(g^n) = nL(g)$.

Suppose we have n independent probability distributions $C^{(i)}$ each with g equally likely events. We can then take an n -fold product (conjunction) and get a distribution with g^n events. Now we apply the third property. We take $A = C^{(1)}$ and $B = C^{(2)} \dots C^{(n)}$ so that $S_I(AB) = L(g^n)$ and $S_I(B) = L(g^{n-1})$. Since A and B are independent, we have $\langle S_I(A|B_{\ell}) \rangle_B = S_I(A) = L(g)$. Thus,

$$L(g) = L(g^n) - L(g^{n-1}),$$

and hence by induction $L(g^n) = nL(g)$.

(e) If $2^m < s^n < 2^{m+1}$, using the results of parts (c) and (d) show

$$\frac{m}{n} < \frac{L(s)}{L(2)} < \frac{m+1}{n}.$$

Show also using the same argument that $m/n < \log s / \log 2 < (m+1)/n$. Hence, show that $|L(s)/L(2) - \log s / \log 2| < 1/n$ and thus $L(s) = k \log s$ for some constant k .

We begin with $2^m < s^n < 2^{m+1}$. Since L is monotonically increasing, we conclude that $L(2^m) < L(s^n) < L(2^{m+1})$. Using the result from (d), we reduce this to $mL(2) < nL(s) < (m+1)L(2)$. We now divide through by $nL(2)$ to find

$$\frac{m}{n} < \frac{L(s)}{L(2)} < \frac{m+1}{n}.$$

Similarly, since \log is also monotonically increasing, $m \log 2 < n \log s < (m+1) \log 2$ so that

$$\frac{m}{n} < \frac{\log s}{\log 2} < \frac{m+1}{n}.$$

Since both rates $L(s)/L(2)$ and $\log(s)/\log(2)$ are confined to the same interval $[m/n, (m+1)/n]$, they must differ by no more than the length of that interval, $1/n$. Given any n no matter how large and any s there is an n with $2^m < s^n < 2^{m+1}$, so we may let $n \rightarrow \infty$, giving

$$L(s) = \frac{L(2)}{\log 2} \log s = k \log s.$$

6.3 Negative Temperature.

(a) **Microcanonical Entropy.** If the net energy is E (corresponding to a number of excited atoms $m = E/\epsilon + N/2$), what is the microcanonical entropy $S_{\text{micro}}(E)$ of our system? Simplify your expression using Stirling's formula, $\log n! = n \log n - n$.

With m atoms in excited states, the microcanonical entropy of the system is

$$\begin{aligned} S_{\text{micro}}(E) &= k_B \log \Omega \\ &= k_B \log \binom{N}{m} = k_B \log \left(\frac{N!}{m!(N-m)!} \right) \\ &\approx k_B [N \log N - m \log m - (N-m) \log(N-m)], \end{aligned}$$

where we have assumed that all quantities are large.

(b) **Negative Temperature.** Find the temperature, using your simplified expression from part (a). (Why is it tricky to do it without approximation?) What happens to the temperature when $E > 0$?

To find the temperature we use

$$\frac{1}{T} = \left. \frac{\partial S}{\partial E} \right|_N = k_B \left[-\frac{\partial m}{\partial E} \log m + \frac{\partial m}{\partial E} \log(N-m) \right],$$

where in terms of energy, the number of excited atoms is $m = \frac{N}{2} + \frac{E}{\epsilon}$. We obtain:

$$T = \frac{\epsilon}{k_B} [\log(N/m - 1)]^{-1} = \frac{\epsilon}{k_B} \left[\log \frac{N/2 - E/\epsilon}{N/2 + E/\epsilon} \right]^{-1}.$$

Note that this would be more tricky without approximations because we would have to take a finite difference of m ! instead of a derivative of m .

For $E > 0$, the argument of the logarithm is less than one, making the temperature $T < 0$.

(c) **Canonical Ensemble: Explicit traces and thermodynamics.** (i) Take one of our atoms and couple it to a heat bath of temperature $k_B T = 1/\beta$. Write explicit formulas for Z_{canon} , E_{canon} , and S_{canon} in the canonical ensemble, as a trace over the two states of the atom.

For one of our atoms coupled to a heat bath at temperature T we have:

$$Z_{\text{canon}} = \text{Tr}(e^{-\beta \mathcal{H}}) = e^{-\beta \epsilon/2} + e^{\beta \epsilon/2} = 2 \cosh\left(\frac{\beta \epsilon}{2}\right),$$

$$E_{\text{canon}} = ((\epsilon/2)e^{-\beta \epsilon/2} - (\epsilon/2)e^{\beta \epsilon/2})/(e^{-\beta \epsilon/2} + e^{\beta \epsilon/2}) = \frac{\text{Tr}(\mathcal{H}e^{-\beta \mathcal{H}})}{Z} = \text{Tr}(\rho \mathcal{H}) = -\frac{\epsilon}{2} \tanh\left(\frac{\beta \epsilon}{2}\right),$$

$$\text{and } S_{\text{canon}} = -k_B \text{Tr}(\rho \log \rho)$$

$$\begin{aligned} &= -k_B \left[\frac{e^{\beta \epsilon/2}}{Z} \log\left(\frac{e^{\beta \epsilon/2}}{Z}\right) + \frac{e^{-\beta \epsilon/2}}{Z} \log\left(\frac{e^{-\beta \epsilon/2}}{Z}\right) \right] \\ &= k_B \log Z + \frac{E}{T}. \end{aligned}$$

(c, ii) Compare the results with what you get by using the thermodynamic relations. Using Z from the trace over states, calculate the Helmholtz free energy A , S as a derivative of A , and E from $A = E - TS$. Do the thermodynamically derived formulas you get agree with the statistical traces?

Using our formula for Z from (i), we calculate A , S , and E as follows:

$$A = -k_B T \log(Z) = -k_B T \log(2 \cosh(\beta \epsilon/2)),$$

$$S = -\left. \frac{\partial A}{\partial T} \right|_{N,V} = \dots = k_B \left[\log Z - \frac{\beta \epsilon}{2} \tanh\left(\frac{\beta \epsilon}{2}\right) \right],$$

$$\text{and } E = A + TS = -\frac{\epsilon}{2} \tanh\left(\frac{\beta \epsilon}{2}\right),$$

in agreement with our results from (i).

(d) What happens to E in the canonical ensemble as $T \rightarrow \infty$? Can you get into the regime discussed in part (b)?

As $T \rightarrow \infty$,

$$E = -\frac{\epsilon}{2} \tanh\left(\frac{\beta\epsilon}{2}\right) \rightarrow 0^-.$$

We can't get into the regime where $E > 0$. Negative temperature states are not obtainable in the canonical ensemble (except perhaps by coupling to a bath at negative temperature).

(e) **Canonical-Microcanonical Correspondence.** Find the entropy in the canonical distribution for N of our atoms coupled to the outside world, from your answer to part (c). How can you understand the value of $S(T = \infty) - S(T = 0)$ simply? Using the approximate form of the entropy from part (b), show that the canonical and microcanonical entropies agree, $S_{\text{micro}}(E) = S_{\text{canon}}(T(E))$. Notice that the two are not equal in the figure above: the form of Stirling's formula we used in part (a) is not very accurate for $N = 50$. In a simple way, explain why the microcanonical entropy is smaller than the canonical entropy.

For N of our atoms coupled to the outside world, but only very weakly to each other (i.e. treat as non-interacting) we have (for distinguishable particles):

$$Z_N = (Z_1)^N.$$

The entropy is therefore extensive:

$$S_N = -\frac{\partial}{\partial T}(-k_B T \log Z_1^N) \Big|_{N,V} = N S_1 = N \left\{ k_B \log Z_1 + \frac{E_1}{T} \right\}.$$

In the limit as $T \rightarrow \infty$ and $T \rightarrow 0$ we find:

$$S(T \rightarrow \infty) = N k_B \log 2,$$

$$\text{and } S(T \rightarrow 0) = 0.$$

As $T \rightarrow \infty$ the thermal energy $k_B T$ becomes much larger than the energy level gap ϵ , and the 2^N states become equally likely; the system tends towards maximum disorder, with $S \rightarrow k_B \log 2^N = N k_B \log 2$ as $T \rightarrow \infty$. As $T \rightarrow 0$ on the other hand, the probability to be in the excited state tends to 0, and the system falls into the unique minimum energy ground state, hence $S = 0$ at $T = 0$.

To show that $S_{\text{micro}}(E) = S_{\text{canon}}(T(E))$, we first invert the formula for energy in the canonical ensemble to obtain $\beta = \beta(E)$:

$$\frac{\beta\epsilon}{2} = -\tanh^{-1}\left(\frac{2E}{\epsilon N}\right).$$

Then we calculate:

$$S_{\text{can}}(T(E)) = N k_B \log \left[2 \cosh(\tanh^{-1}(2E/\epsilon N)) \right] - 2 \frac{k_B E}{\epsilon} \tanh^{-1}(2E/\epsilon N).$$

Using the relation for $\tanh^{-1} x$ given in the problem, we let $x \equiv 2E/\epsilon N$ and find:

$$S_{\text{can}}(T(E)) = N k_B \left[\log 2 - \frac{1}{2}(1+x) \log(1+x) - \frac{1}{2}(1-x) \log(1-x) \right].$$

Noting that $1+x = 1 + \frac{2E}{\epsilon N} = \frac{2m}{N}$ and $1-x = \frac{2(N-m)}{N}$, we find:

$$S_{\text{can}}(T(E)) = \dots = k_B \left\{ N \log N - m \log m - (N-m) \log(N-m) \right\} = S_{\text{micro}}(E).$$

The entropies are equal here because we assumed very large N and $N - m$. In general the canonical entropy is greater than the microcanonical entropy because we have less information about the state of a system in the canonical ensemble than in the micro-canonical. Namely, in the micro-canonical our system is confined to an energy surface, while in the canonical it is not—we average over all the energy surfaces. For very large number of particles, the equilibrium energy is by far the most likely, so the average over all the energy surfaces weighted by their probabilities approaches the average over just the equilibrium energy surface.

(f) **Fluctuations.** Calculate the root-mean-square energy fluctuations for N of our atoms. Evaluate it at $T(E)$ from part (b): it should have a particularly simple form. For large N , are the fluctuations in E small compared to E ?

Start from $E = -\frac{\partial}{\partial \beta} \log Z$:

$$\frac{\partial E}{\partial \beta} = \frac{\partial T}{\partial \beta} \frac{\partial E}{\partial T} = \frac{1}{Z^2} \left(\frac{\partial Z}{\partial \beta} \right)^2 - \frac{1}{Z} \frac{\partial^2 Z}{\partial \beta^2} = \langle E \rangle^2 - \langle E^2 \rangle = \sigma_E^2.$$

For N of our atoms, we find:

$$\sigma_E = \frac{\epsilon}{2} \sqrt{N} \operatorname{sech} \left(\frac{\beta \epsilon}{2} \right) = \epsilon \sqrt{N} \frac{1}{e^{\beta \epsilon/2} + e^{-\beta \epsilon/2}}$$

Evaluating this at $T(E) = \epsilon / (k_B \log(N/m - 1))$ from part (b), $\beta \epsilon / 2 = \frac{1}{2} \log(N/m - 1)$, and we find

$$\sigma_E = \epsilon \sqrt{N} \frac{1}{\sqrt{(N-m)/m} + \sqrt{m/(N-m)}} = \epsilon \sqrt{N} \frac{\sqrt{m(N-m)}}{\sqrt{(N-m)^2 + m^2}} = \frac{\epsilon}{\sqrt{N}} [m(N-m)]^{1/2}.$$

At moderate temperatures, $m \ll N$ and $E = (-N/2 + m)\epsilon \approx -N\epsilon/2$. Then:

$$\frac{\sigma_E}{E} \rightarrow \frac{2\sqrt{m}}{N},$$

which is very small for large N . (At very high temperatures such that $m \rightarrow N/2$, $E \rightarrow 0$, so the fractional fluctuations $\sigma_E/E \rightarrow \infty$.)

6.4 Molecular Motors: Which Free Energy?.

(a) Without knowing anything further about the chemistry or biology of the system, which two of the following must be true on average, in all cases, according to basic laws of thermodynamics?

(1) The total entropy of the universe must increase or stay unchanged with time.

True. This is the second law of thermodynamics.

(2) The entropy S_s of the system must increase with time.

False. The entropy of the system may decrease or increase, so long as (a) is true.

(3) The total energy E_T of the universe must decrease with time.

False. Total energy is conserved.

(4) The energy E_s of the system must decrease with time.

False. It is the appropriate free energy that must decrease.

(5) $G_s - Fx = E_s - TS + PV_s - Fx$ must decrease with time, where G_s is the Gibbs free energy of the system.

True. $G_s - Fx$ is the relevant free energy for this problem. It's important that the $-Fx$ term not be left out. The free energy represents the energy available for the system to do work on the outside world, and there is potential energy associated with the constant force F .

To see this last point explicitly, and in particular to verify the minus sign multiplying Fx , is a bit tricky to do formally. (It's straightforward to derive from physical considerations.) In full detail, it can be seen as follows:

Let's say the universe, which we'll consider from a microcanonical point of view, is composed of our system and the bath it sits in. The system can exchange heat, volume and distance with the external bath, so the thermodynamic identity reads:

$$dE_s = TdS_s - p dV_s + F dx$$

where F is the force exerted on the system by the bath. Clearly, $F dx$ is thus the work done on the system by the bath for a displacement dx of the system.

At equilibrium, the total entropy of the universe must be maximized. It is given by:

$$\begin{aligned} S_{\text{universe}}(E, V, L) &= \sum_{\{E_s^*, V_s^*, x^*\}} (S_s(E_s^*, V_s^*, x^*) + S_{\text{bath}}(E - E_s^*, V - V_s^*, L - x^*)) \\ &\simeq S_s(E_s, V_s, x) + S_{\text{bath}}(E - E_s, V - V_s, L - x) \end{aligned}$$

where it is understood that the configuration $\{E_s, V_s, x\}$ is the configuration that maximizes the summand. Next, we want to Taylor expand the bath entropy term, since we expect $\{E_s, V_s, x\}$ to be a very small perturbation about $\{E, V, L\}$. We already know that $\frac{\partial S}{\partial E} = \frac{1}{T}$ and $\frac{\partial S}{\partial V} = \frac{p}{T}$; what about $\frac{\partial S}{\partial x}$? From eqn 3.33,

$$\begin{aligned} \left. \frac{\partial f}{\partial x} \right|_y \left. \frac{\partial x}{\partial y} \right|_f \left. \frac{\partial y}{\partial f} \right|_x &= -1 \\ \Rightarrow \left. \frac{\partial S}{\partial x} \right|_E \left. \frac{\partial x}{\partial E} \right|_S \left. \frac{\partial E}{\partial S} \right|_x &= -1 \\ \left. \frac{\partial S}{\partial x} \right|_E &= \frac{-F}{T} \end{aligned}$$

where the thermodynamic identity has been used in the last equation. We can thus write an expression for the entropy of the universe as a Taylor expansion:

$$\begin{aligned} S_{\text{universe}} &\simeq S_s + S_{\text{bath}}(E, V, L) + \left[-E_s \frac{\partial S_{\text{bath}}}{\partial E} - V_s \frac{\partial S_{\text{bath}}}{\partial V} - x \frac{\partial S_{\text{bath}}}{\partial x} \right] \\ &= S_{\text{bath}}^{(0)} - \frac{1}{T} [E_s - TS_s + pV_s - Fx] = S_{\text{bath}}^{(0)} - \frac{1}{T} (G_s - Fx) \end{aligned}$$

For the entropy of the universe to be maximized, we see that $G_s - Fx$ must be minimized. Hence, if we start from a non equilibrium situation, we see that $G_s - Fx$ will decrease with time as the universe equilibrates.

(b) Of the following statements, which are true, assuming that the pulled RNA is in equilibrium?

- (1) $\frac{\rho_z}{\rho_u} = \exp((S_z^{\text{tot}} - S_u^{\text{tot}})/k_B)$
- (2) $\frac{\rho_z}{\rho_u} = \exp((E_z - E_u)/k_B T)$
- (3) $\frac{\rho_z}{\rho_u} = \exp((G_z - G_u)/k_B T)$
- (4) $\frac{\rho_z}{\rho_u} = \exp((G_z - G_u + FL)/k_B T)$

We can simply apply results from the previous part of the question.

- (1) True.
- (2) False.
- (3) False.
- (4) True.

6.5 Laplace.

Show that the canonical partition function $Z(\beta)$ can be written as the Laplace transform of the microcanonical volume of the energy shell $\Omega(E)$.

The microcanonical partition function is the volume of the energy surface in phase space

$$\Omega_E = \int d\mathbb{V} d\mathbb{P} \delta(E - \mathcal{H}(\mathbb{V}, \mathbb{P})).$$

The canonical partition function is

$$Z_{\text{can}} = \int d\mathbb{V} d\mathbb{P} e^{-(\beta \mathcal{H}(\mathbb{V}, \mathbb{P}))}.$$

Inserting $\int_0^\infty dE \delta(E - \mathcal{H}(\mathbb{V}, \mathbb{P})) = 1$ (assuming $E > 0$) into the formula for Z_{can} and interchanging the order of integration, we get

$$Z_{\text{can}} = \int dE \int d\mathbb{V} d\mathbb{P} \delta(E - \mathcal{H}(\mathbb{V}, \mathbb{P})) e^{-(\beta \mathcal{H}(\mathbb{V}, \mathbb{P}))} = \int_0^\infty dE e^{-\beta E} \Omega_E = \mathcal{L}\{\Omega_E\}(\beta),$$

the Laplace transform of the microcanonical partition function.

6.7 Legendre.

Using your Laplace transform of exercise 6.5, find an equation for E_{max} where the integrand is maximized. Does this energy equal the energy which minimizes the Legendre transform 6.73? Approximate $Z(\beta)$ in your Laplace transform by the value of the integrand at this maximum (ignoring fluctuations). Does it give the Legendre transform relation 6.73?

Yes, they both agree.

Laplace transform: The partition function in the canonical ensemble is written in terms of the microcanonical energy-shell volume $\Omega(E)$ as the Laplace transform $Z = \int e^{-\beta E} \Omega(E) dE$. The integrand is maximized when $\partial[e^{-\beta E} \Omega(E)]/\partial E = -\beta e^{-\beta E} \Omega(E) + e^{-\beta E} \partial \Omega / \partial E = 0$, so $\beta = 1/(k_B T) = (1/\Omega(E)) \partial \Omega / \partial E = \partial \log(\Omega) / \partial E$. But since the microcanonical entropy is $S = k_B \log \Omega$, the peak probability E_{max} occurs where $1/T = \partial S / \partial E$ (as usual).

Legendre transform: Here we must minimize $E(S) - TS$ with respect to S . But this tells us that $\partial E/\partial S = -T$. Since both the canonical and microcanonical ensembles hold V and N fixed, $\partial E/\partial S = 1/(\partial S/\partial E)$, and the energy at the entropy which minimizes the Legendre transform equals the energy which maximizes the integrand in the Laplace transform. Finally, if we approximate the Laplace transform giving Z by the value of its integrand at the latter's maximum, $Z \approx e^{-\beta E_{\max}} \Omega(E_{\max}) = e^{-\beta E_{\max}} e^{S/k_B}$, we find that $A = -k_B T \log Z = E - TS$, the Legendre transform.

6.8 Euler.

(a) Using the fact that entropy $S(N, V, E)$ is extensive, show that

$$N \left. \frac{\partial S}{\partial N} \right|_{V,E} + V \left. \frac{\partial S}{\partial V} \right|_{N,E} + E \left. \frac{\partial S}{\partial E} \right|_{N,V} = S.$$

Show from this that in general $S = (E + pV - \mu N)/T$ and hence $E = TS - pV + \mu N$.

The entropy is extensive, meaning if we increase the size of our system by some factor α , then the entropy increases by the same factor:

$$S(\alpha N, \alpha V, \alpha E) = \alpha S(N, V, E).$$

Now let $\alpha = 1 + \epsilon$, add and subtract some terms on the left, and obtain:

$$\begin{aligned} N \cdot \frac{(S(\alpha N, \alpha V, \alpha E) - S(N, \alpha V, \alpha E))}{\epsilon N} + V \cdot \frac{(S(N, \alpha V, \alpha E) - S(N, V, \alpha E))}{\epsilon V} \\ + E \cdot \frac{(S(N, V, \alpha E) - S(N, V, E))}{\epsilon E} = S(N, V, E). \end{aligned}$$

Taking the limit as $\epsilon \rightarrow 0$ yields the desired expression:

$$N \left. \frac{\partial S}{\partial N} \right|_{V,E} + V \left. \frac{\partial S}{\partial V} \right|_{N,E} + E \left. \frac{\partial S}{\partial E} \right|_{N,V} = S(N, V, E). \quad (1)$$

Recall that

$$\frac{P}{T} = \left. \frac{\partial S}{\partial V} \right|_{N,E}, \quad -\frac{\mu}{T} = \left. \frac{\partial S}{\partial N} \right|_{E,V}, \quad \text{and} \quad \frac{1}{T} = \left. \frac{\partial S}{\partial E} \right|_{N,V}.$$

Plugging these into (1) leads directly to Euler's equation:

$$E = TS - PV + \mu N.$$

(b) Test this explicitly for the ideal gas. Use the ideal gas entropy (eqn 3.57)

$$S(N, V, E) = \frac{5}{2} N k_B + N k_B \log \left[\frac{V}{N h^3} \left(\frac{4\pi m E}{3N} \right)^{3/2} \right],$$

to derive formulae for T , P , and μ in terms of E , N , and V , and verify eqn 6.75.

For the ideal gas, we showed in problem 2.4(c) that

$$T = \frac{2E}{3Nk_B}, \quad P = \frac{2E}{3V}, \quad \text{and} \quad \mu = -k_B T \log \left[\frac{V}{N h^3} \left(\frac{4\pi m E}{3N} \right)^{3/2} \right].$$

Substituting these into Euler's equation gives

$$\begin{aligned}
 S &= \frac{1}{T}(E + PV - \mu N) \\
 &= \frac{3k_B N}{2E} \left(E + \frac{2E}{3V} V + k_B N \frac{2E}{3k_B N} \log \left[\frac{V}{N h^3} \left(\frac{4\pi m E}{3N} \right)^{\frac{3}{2}} \right] \right) \\
 &= k_B N \left\{ \frac{5}{2} + \log \left[\frac{V}{N h^3} \left(\frac{4\pi m E}{3N} \right)^{\frac{3}{2}} \right] \right\}.
 \end{aligned}$$

agreeing with eqn 3.57.

6.9 Gibbs-Duhem.

(a) Write μ as a suitable derivative of the Gibbs free energy.

Given $G = E - TS + PV$ and the differential form $dE = T dS - P dV + \mu dN$,

$$dG = dE - d(TS) + d(PV) = -S dT + V dP + \mu dN,$$

giving us

$$\mu = \left. \frac{\partial G}{\partial N} \right|_{T,P}$$

(b) Argue that changing the number of particles in a large system at fixed temperature and pressure should not change the chemical potential.

In words, adding an extra particle at fixed P and T to a large, uniform, equilibrium system should not cost a different free energy if we double the size of the system. (This will not be true of nonextensive systems like gravitationally bound groups of stars!)

We can also derive this symbolically. Since G is intensive, $G(\lambda N) = \lambda G(N)$, and hence $\mu(\lambda N) = (G(\lambda N + \epsilon) - G(\lambda N))/\epsilon = \lambda(G(N + \epsilon/\lambda) - G(N))/\epsilon = \lambda(1/\lambda)(\partial G/\partial N)|_N = \mu(N)$. and thus μ is independent of λ , and thus of N .

(c) Use the Euler relation to write a formula for G . Is it indeed proportional to N ? What about your formula for μ from part (a); will it be dependent on N ?

Using the Euler relation $E = TS - PV + \mu N$ we derive $G = E - TS + PV = \mu N$. It is proportional to N only if μ is independent of N , a result we derived in part (b). Similarly, the formula from part (a) $\mu = (\partial G/\partial N)|_{T,P}$ is independent of N only if $(\partial^2 G/\partial N^2)|_{T,P} = 0$, another way of looking at the consequence of extensivity, part (b).

6.10 Clausius-Clapeyron.

(a) Apply the Gibbs-Duhem relation to both phases, for a small shift by ΔT along the phase boundary. Let $s_1, v_1, s_2,$ and v_2 be the molecular entropies and volumes; derive the Clausius-Clapeyron equation for the slope of the coexistence line on the phase diagram

$$\frac{dP}{dT} = (s_1 - s_2)/(v_1 - v_2).$$

Applying the Gibbs-Duhem relation to both phases for a small shift along the phase boundary, we find

$$0 = S_1 dT_1 - V_1 dP_1 + N_1 d\mu_1$$

$$\text{and } 0 = S_2 dT_2 - V_2 dP_2 + N_2 d\mu_2.$$

Dividing out the N_1 and N_2 , and noting since the phases are in equilibrium along the boundary that $dT_1 = dT_2$, $dP_1 = dP_2$, and $d\mu_1 = d\mu_2$, we find the Clausius-Clapeyron equation:

$$s_1 dT - v_1 dP = s_2 dT - v_2 dP$$

$$\iff$$

$$(s_1 - s_2)dT = (v_1 - v_2)dP$$

$$\iff$$

$$\frac{dP}{dT} = \frac{(s_1 - s_2)}{(v_1 - v_2)} \text{ along the co-existence curve.}$$

(b) Write a formula for dP/dT not involving the entropy.

The entropy difference per molecule upon a phase transition is

$$\Delta s = \frac{Q/N}{T} = \frac{L}{T} = s_1 - s_2,$$

where L is the latent heat per molecule. Therefore, we have a formula for dP/dT not involving the entropy:

$$\frac{dP}{dT} = \frac{L/T}{(v_1 - v_2)}.$$

6.11 Barrier Crossing.

(a) Let the probability that a particle has position X be $\rho(X)$. What is the ratio of probability densities $\rho(X_B)/\rho(X_0)$ if the particles near the top of the barrier are assumed in equilibrium with those deep inside the well?

The particles near the top of the barrier are assumed to be in equilibrium with those near the bottom of well. This means that they are at the same temperature, so we may write

$$\frac{\rho(X_B)}{\rho(X_0)} = \frac{\exp(-E_B/k_B T)}{\exp(-E_0/k_B T)} = \exp(-B/k_B T),$$

since $E_0 = 0$ and $E_B = B$.

As a side note, because the temperature is the same, the expected velocities are also the same for particles whether they are at the top or bottom of the potential!

(b) In this approximation, what is the probability density $\rho(X)$ near the bottom of the well? What is $\rho(X_0)$, the probability density of having the system at the bottom of the well?

In the approximation that $U(X)$ is quadratic, $\rho(X) \propto \exp(-U(X)/k_B T)$ near the bottom of the well. Normalization of the Gaussian determines the constant of proportionality:

$$\rho(X) = \sqrt{\frac{m\omega^2}{2\pi k_B T}} \exp\left(-\frac{m\omega^2}{2k_B T}(X - X_0)^2\right)$$

$$\text{and } \rho(X_0) = \sqrt{\frac{m\omega^2}{2\pi k_B T}}.$$

(c) First give a formula for the decay rate Γ , for an unknown probability density $\rho(X_B)\rho(V)$ as an integral over the velocity V . Then, using your formulas from parts (a) and (b), give your estimate of the decay rate for our system.

Particles near the top of the barrier that are moving to the right with speed V will cross the barrier in time Δt if they are within a distance $V\Delta t$ from the top. Therefore $\Gamma_V = \rho(X_B)(V\Delta t)/\Delta t = \rho(X_B)V$ is the probability per unit time of crossing the barrier to the right for particles moving with speed V . The total rate is therefore

$$\begin{aligned} \Gamma &= \int_0^\infty \Gamma_V \rho(V) dV = \rho(X_B) \int_0^\infty V \rho(V) dV \\ &= \rho(X_0) \exp(-B/k_B T) \int_0^\infty V \rho(V) dV. \end{aligned}$$

Using the probability distribution $\rho(V)$ given in the problem we obtain:

$$\Gamma = \frac{\omega}{2\pi} e^{-B/k_B T}.$$

The general, multidimensional transition-state theory rate formula [55] is

$$\Gamma = \frac{\prod_i \omega_i}{2\pi \prod_j \omega_j^B} \exp(-B/k_B T)$$

where ω_i are the eigenfrequencies of the well and ω_j^B are the eigenfrequencies at the saddlepoint forming the barrier, with the unstable normal mode omitted.

6.13 Pollen and Hard Squares.

(a) What is the area $A(Q \gg 0)$ available for the gas molecules, in units of $(\text{length})^2$, when the pollen grain is far from the wall? What is the overlap of the excluded regions, $A(0) - A(\infty)$, when the pollen grain touches the wall, $Q = 0$? Give formulas for $A(Q)$ as a function of Q for the two relevant regions, $Q < b$ and $Q > b$.

The area available to the gas molecules when the pollen grain is far from the wall is:

$$A(Q \gg 0) = (L - b)^2 - (B + b)^2.$$

When the grain is touching the wall (far from a corner):

$$A(Q = 0) = (L - b)^2 - (B + b)B,$$

so the overlap of the excluded regions is:

$$A(0) - A(\infty) = b(B + b).$$

For arbitrary Q , the area is:

$$A(Q) = (L - b)^2 - (B + b)^2 + (b - Q)(B + b) \Theta(b - Q),$$

where $\Theta(x)$ is the step function.

(b) What is the configuration-space volume $\Omega(Q)$ for the gas, in units of $(\text{length})^{2N}$? What is the configurational entropy of the ideal gas, $S(Q)$?

Assuming no interactions between the gas molecules, the configuration space volume for the gas is:

$$\Omega(Q) = \frac{[A(Q)]^N}{N!}.$$

The corresponding configurational entropy is then:

$$S = k_B \log \Omega(Q) = k_B (N \log A(Q) - \log N!) \approx k_B [N \log A(Q) - (N \log N - N)] \approx k_B N (\log(A(Q)/N) + 1).$$

(c) What is the resulting coarse-grained free-energy of the pollen grain, $\mathcal{F}(Q) = E - TS(Q)$, in the two regions $Q > b$ and $Q < b$? Use $\mathcal{F}(Q)$ to calculate the force on the pollen grain for $Q < b$. Is the force positive (away from the wall) or negative? Why?

Since there are no interactions, the energy E is purely kinetic and the free energy is:

$$\mathcal{F}(Q) = E_{\text{kinetic}} - TS(Q).$$

The force on the pollen grain for $Q < b$ is therefore:

$$F = -\frac{\partial \mathcal{F}}{\partial Q} = T \frac{\partial S}{\partial Q} = \dots = -\frac{Nk_B T}{A(Q)} (B + b) \Theta(b - Q).$$

So there is an effective force on the pollen grain pushing it towards the wall—a consequence of the tendency for entropy to increase, or equivalently for the free energy to decrease.

(d) Directly calculate the force due to the ideal-gas pressure on the far side of the pollen grain, in terms of $A(Q)$. Compare it to the force from the partial trace in part (c). Why is there no balancing force from the other side? Effectively how “long” is the far side of the pollen grain?

In two dimensions, the ideal gas law "PV" – $PA(Q) = Nk_B T$ gives the force on the far side as:

$$F = -(\text{pressure}) \times (\text{length}) = -PB = -\frac{Nk_B T}{A(Q)} B.$$

This is the same as in part (c) except for a correction since the gas molecules whose centers hit within $b/2$ of the corners also exert pressure. There is no balancing force from the near side because when $Q < b$ the gas molecules can't fit between the grain and the wall. From the force we calculated in part (c), we see that the effective length of the side is $B + b$.

6.14 Statistical Mechanics and Statistics.

(a) True or false: This probability density corresponds...

True, because with H and $k_B T$ defined as in the problem the probability is

$$P(D|M) \propto \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^N (y^{(M)}(t_i) - y_i)^2\right) = \exp(-H/k_B T).$$

(b) Prove Bayes' theorem ...

$P(M \text{ and } D) = P(M|D)P(D)$. But also $P(M \text{ and } D) = P(D|M)P(M)$. Hence $P(M|D)P(D) = P(D|M)P(M)$, so $P(M|D) = P(D|M)P(M)/P(D)$.

(c) Which contour plot...

C. Because:

$$\begin{aligned} P(M|D) &= P(M) P(D|M) \\ &= \frac{1}{4} \prod_{i=1}^2 \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(y^{(M)}(t_i) - y_i)^2}{2\sigma^2}\right] \\ &= \frac{1}{8\pi\sigma^2} \exp\left[-\frac{1}{2\sigma^2} \left\{ (b-1)^2 + (m+b-2)^2 \right\}\right] \Rightarrow C. \end{aligned}$$

7.1 Ensembles and quantum statistics.

This exercise is not easy, but it is rewarding. The problem demands both physical reasoning and formal calculations, and should provide some intuition about quantum statistics, the relationships between the three main statistical ensembles (microcanonical, canonical, and grand canonical), and (in part (c)) chemical potentials.

(a) The system is first held at constant energy. In figure 7.12 which curve represents the entropy of the fermions as a function of the energy? Bosons? Distinguishable particles? Maxwell-Boltzmann particles?

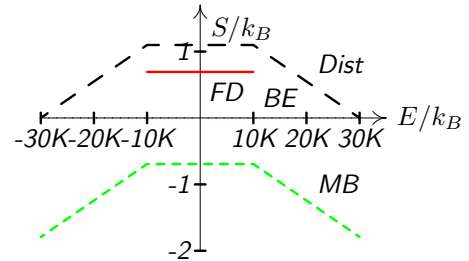
An isolated system at constant energy follows the microcanonical ensemble. The entropy is therefore $S = k_B \log \Omega(E)$.

Fermions. We have spin-1/2 fermions, so that only two particles are allowed in the same energy level. Thus, the states $|++\rangle$ and $|--\rangle$ are disallowed, and we therefore have only allowed energy level configurations $|+-\rangle$ and $|+--\rangle$, with energies ϵ and $-\epsilon$ respectively. Since the particles are indistinguishable, the order is unimportant. But each of these states does have a two-fold degeneracy since the unpaired level can have either of the two allowed spin states. Thus, $\Omega_{FD}(\pm 3\epsilon) = 0$ and $\Omega_{FD}(\pm \epsilon) = 2$. Thus, $S_{FD}(\pm 3\epsilon)$ is undefined, while $S_{FD}(\pm \epsilon) = k_B \log 2 \Rightarrow \mathbf{D}$.

Bosons. For bosons, there are exactly four possible states: $|+++ \rangle$, $|++-\rangle$, $|+-\rangle$, and $|--\rangle$ with energies 3ϵ , ϵ , $-\epsilon$, and -3ϵ respectively. Thus, each energy has $\Omega_{BE}(E) = 1$ so that for all possible E , $S_{BE}(E) = 0 \Rightarrow \mathbf{B}$.

Distinguishable. For distinguishable particles, there is one state each with energy $\pm 3\epsilon$ and three permutations each with energy $\pm \epsilon$. Thus, $\Omega_D(\pm 3\epsilon) = 1$ and $\Omega_D(\pm \epsilon) = 3$, so that $S_D(\pm 3\epsilon) = 0$ and $S_D(\pm \epsilon) = \log 3 \Rightarrow \mathbf{C}$.

Maxwell-Boltzmann. This is the same as distinguishable, except we get a Gibbs factor of $\Omega_{MB} = \Omega_D/N!$, so that for $N = 3$ we get $S_{MB} = S_D - k_B \log 6 \Rightarrow \mathbf{E}$.



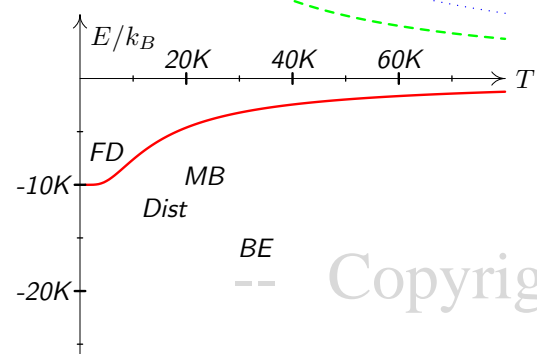
(b) The system is now held at constant temperature. In figure 7.13 which curve represents the mean energy of the fermions as a function of temperature? Bosons? Distinguishable particles? Maxwell-Boltzmann particles?

Let us start by trying to deduce which curve corresponds to which kind of statistics without detailed calculations. We note that the ground-state energy for the Fermions is $-10K$, and for the other three systems it is $-30K$, so we can tell without calculations that A is the energy curve for Fermi-Dirac statistics. Also, at high temperatures all states are equally weighted, so $\langle E \rangle = 0$ for all four types of statistics—ruling out curve D and perhaps curve B (depending upon where the latter asymptotes to). Now, Distinguishable and Maxwell-Boltzmann differ only in their normalization of Z , so their energies $E = \partial(\log Z)/\partial\beta$ will have the same temperature dependence. Also, these two have three low-lying excited states with $E/k_B = 10K$, where Bose-Einstein has only one excited state, so we expect their energies to rise faster than Bose-Einstein as the temperature is raised. This suggests that curve C corresponds to both Maxwell-Boltzmann and Distinguishable, and E corresponds to Bose-Einstein.

Having explored qualitative arguments, let us now do the calculations explicitly. Holding temperature fixed implies the presence of a heat bath, so we are in the canonical ensemble. We can now write the partition function and find $\langle E \rangle = -\frac{\partial}{\partial\beta} \log Z$.

Fermions.

$$Z_{FD} = \sum_k \Omega_{FD}(\epsilon_k) e^{-\beta \epsilon_k} = \cosh(\beta \epsilon),$$



where the sum is over only the $\pm\epsilon$ states since $\Omega_{\text{FD}}(\pm 3\epsilon) = 0$.

$$\langle E \rangle_{\text{FD}} = -\epsilon \tanh(\beta\epsilon) \quad \Rightarrow \mathbf{A}.$$

Bosons.

$$Z_{\text{BE}} = \sum_k \Omega_{\text{BE}}(\epsilon_k) e^{-\beta\epsilon_k} = 2 \cosh(3\beta\epsilon) + 2 \cosh(\beta\epsilon).$$

$$\langle E \rangle_{\text{BE}} = -\epsilon \frac{3 \sinh(3\beta\epsilon) + \sinh(\beta\epsilon)}{\cosh(3\beta\epsilon) + \cosh(\beta\epsilon)} \quad \Rightarrow \mathbf{E}.$$

Distinguishable.

$$Z_{\text{D}} = \sum_k \Omega_{\text{D}}(\epsilon_k) e^{-\beta\epsilon_k} = 2 \cosh(3\beta\epsilon) + 6 \cosh(\beta\epsilon).$$

$$\langle E \rangle_{\text{D}} = -3\epsilon \frac{\sinh(3\beta\epsilon) + \sinh(\beta\epsilon)}{\cosh(3\beta\epsilon) + 3 \cosh(\beta\epsilon)} \quad \Rightarrow \mathbf{C}.$$

Maxwell-Boltzmann.

$$Z_{\text{MB}} = \sum_k \Omega_{\text{MB}}(\epsilon_k) e^{-\beta\epsilon_k} = \frac{Z_{\text{D}}}{6}.$$

We note that this factor will not matter since $\frac{\partial}{\partial \beta} \log Z_{\text{MB}} = \frac{\partial}{\partial \beta} (\log Z_{\text{D}} - \log 6)$.

$$\langle E \rangle_{\text{MB}} = \langle E \rangle_{\text{D}} \quad \Rightarrow \mathbf{C}.$$

(c) The system is now held at constant temperature, with chemical potential set to hold the average number of particles equal to three. In figure 7.14, which curve represents the chemical potential of the fermions as a function of temperature? Bosons? Distinguishable? Maxwell-Boltzmann?

Here it is not straightforward to deduce the answers purely from qualitative arguments. However, the analytical calculations will often have two formal solutions, only one of which is physically relevant. To distinguish the correct solution, we will use the behavior at low temperatures.

The chemical potential at low temperatures represents the energy needed to add or subtract a particle to the system. (If the energy to add and to subtract differ from one another, then the chemical potential lies between the two energies.) For Fermions, the state at $E/k_B = +10$ has one particle, so $\mu_F \rightarrow 10k_B$ as $T \rightarrow 0$. For the other three distributions, a new particle will be added to the lower state at low temperatures, so $\mu \rightarrow -10k_B$ as $T \rightarrow 0$. Also, for Bosons the $\mu(T)$ curve must lie below $-10k_B$ for all temperatures.

We now turn to calculating each curve. Now that we have both a heat bath and a particle bath, we must use the grand canonical ensemble. If we put T and μ into this, we can determine $\langle E \rangle$ and $\langle n \rangle$. But we are given $\langle n \rangle = 3$, so we must work backwards to find μ that satisfies this. In general can factor Ξ in two equivalent ways,

$$\Xi = \sum_n Z_n e^{\beta\mu n} = \prod_k \Xi_k,$$

where Z_n is the appropriate n -particle partition function and Ξ_k is the appropriate grand partition function for just the k th mode. For all the indistinguishable particles, the results are already derived for us in terms of $\langle n_k \rangle$. We will need to start from scratch for distinguishable, and we will also go into a little more detail for Maxwell-Boltzmann to contrast it with the

distinguishable particles.

Fermions. We have the average occupancy of a single-particle state k is given by

$$\langle n_k \rangle_{\text{FD}} = \frac{1}{e^{\beta(\epsilon_k - \mu)} + 1}.$$

We have four single-particle states: two levels each with two spins. Thus,

$$3 = \sum_k \langle n_k \rangle_{\text{FD}} = \frac{2}{e^{\beta(\epsilon - \mu)} + 1} + \frac{2}{e^{-\beta(\epsilon + \mu)} + 1}.$$

We solve for $e^{-\beta\mu}$ and get two solutions. Throwing out the negative solution (which goes to the wrong value as $T \rightarrow 0$) yields

$$\mu_{\text{FD}} = -\frac{1}{\beta} \log \left(\frac{1}{3} \sqrt{\cosh^2(\beta\epsilon) + 3} - \frac{1}{3} \cosh(\beta\epsilon) \right) \Rightarrow \mathbf{C}.$$

We could also expand this to find the linear behavior at large T of curve C.

Bosons. Again we use the equation from the text,

$$\langle n_k \rangle_{\text{BE}} = \frac{1}{e^{\beta(\epsilon_k - \mu)} - 1}.$$

Now we have only the two states, so we add their average occupations.

$$3 = \sum_k \langle n_k \rangle_{\text{BE}} = \frac{1}{e^{\beta(\epsilon - \mu)} - 1} + \frac{1}{e^{-\beta(\epsilon + \mu)} - 1}.$$

Solving again for $e^{-\beta\mu}$ we again have two solutions; we choose the one with $\mu \rightarrow -\epsilon$ as $T \rightarrow 0$. This gives

$$\mu_{\text{BE}} = -\frac{1}{\beta} \log \left(\frac{1}{3} \sqrt{16 \cosh^2(\beta\epsilon) - 15} + \frac{4}{3} \cosh(\beta\epsilon) \right) \Rightarrow \mathbf{E}.$$

Distinguishable. Distinguishable particles are tricky, since we don't have $\langle n_k \rangle_{\text{D}}$ given to us. We write the total grand partition function

$$\Xi_{\text{D}} = \sum_n \left(\sum_k e^{-\beta(\epsilon_k - \mu)} \right)^n = \sum_n 2e^{n\beta\mu} \cosh^n(\beta\epsilon) = \frac{1}{1 - 2e^{\beta\mu} \cosh(\beta\epsilon)},$$

where we assume convergence in the geometric series (we'll check later). Notice that if we expand the power $(\cdot)^n$, in the first expression for Ξ , we end up with appropriate counting for distinguishable particles. We now write the grand free energy $\Phi_{\text{D}} = -k_B T \log \Xi_{\text{D}}$ and differentiate

$$\langle n \rangle_{\text{D}} = -\frac{\partial \Phi}{\partial \mu} = \frac{k_B T}{\Xi} \frac{\partial \Xi}{\partial \mu} = 3.$$

Solving this gives

$$\mu_{\text{D}} = -\frac{1}{\beta} \log \left(\frac{8}{3} \cosh(\beta\epsilon) \right) \Rightarrow \mathbf{F}.$$

We can now check self-consistently that the series above converges since $2e^{\beta\mu} \cosh(\beta\epsilon) < 1$.

Maxwell-Boltzmann. This is a little more straightforward. We have the grand partition function for each mode

$$\Xi_k = \sum_{n=0}^{\infty} \frac{1}{n!} \left(e^{-\beta(\epsilon_k - \mu)} \right)^n.$$

Now the total grand partition function is the product of the grand partition function for each mode.

The curious reader may verify that multiplying $\prod_k \Xi_k$ as it is written above results in counting every configuration exactly once, ignoring permutations (thanks to the $1/n!$). This is precisely what we need for Maxwell-Boltzmann particles. The book leads us in §7.5 to

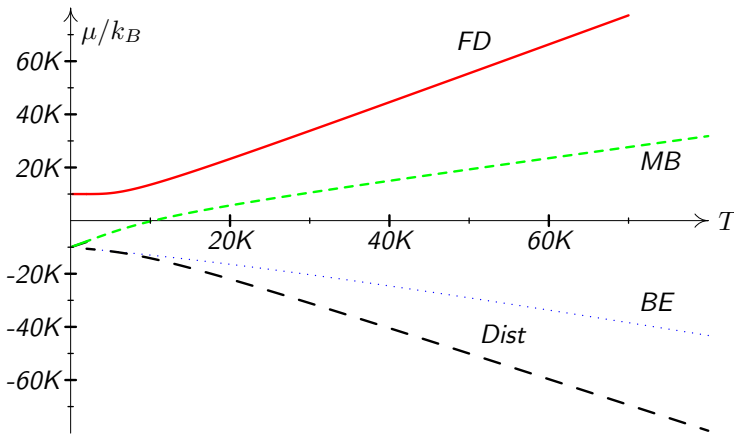
$$\langle n_k \rangle_{\text{MB}} = e^{-\beta(\epsilon_k - \mu)}.$$

Adding up the average occupation of both modes,

$$3 = \sum_k \langle n_k \rangle_{\text{MB}} = e^{\beta\mu} (e^{-\beta\epsilon} + e^{\beta\epsilon}) = 2e^{\beta\mu} \cosh(\beta\epsilon).$$

We can solve this easily,

$$\mu_{\text{MB}} = -\frac{1}{\beta} \log(2 \cosh(\beta\epsilon)/3) \Rightarrow \mathbf{A}.$$



	Fermi	Bose	Dist.	MB
(a)	D	B	C	E
(b)	A	E	C	C
(c)	C	E	F	A

7.2 Phonons and Photons are Bosons.

(a) Show that the ...

The canonical partition function for a quantum harmonic oscillator of frequency ω is:

$$Z = \sum_{n=0}^{\infty} e^{-\beta\epsilon_n} = \sum_{n=0}^{\infty} e^{-\beta\hbar\omega(n+1/2)} = \frac{e^{-\beta\hbar\omega/2}}{1 - e^{-\beta\hbar\omega}}.$$

For bosons multiply filling a single state of energy $\hbar\omega$ with $\mu = 0$, the grand canonical partition function is:

$$\Xi = \sum_{n=0}^{\infty} e^{-\beta\hbar\omega n} = \frac{1}{1 - e^{-\beta\hbar\omega}},$$

which is the same as the canonical, except for a shift of the total (average) energy by $\hbar\omega/2$.

(b) Do phonons or photons ...

Phonons and photons do not Bose condense at low temperatures, because their number is not conserved. The fact that $\mu = 0$ and $\epsilon_{\min} > 0$ means the average occupancy of all the modes goes to zero as $T \rightarrow 0$. Here it may be more appropriate

to think of phonons and photons as excitations of harmonic modes rather than as particles. Note however that one can have a Bose condensation of photons in a non-thermal distribution. A laser is one example of a macroscopic number of photons all in the same harmonic mode.

7.3 Phase Space Units and the Zero of Entropy.

(a) **Arbitrary zero of the classical entropy.** Show that the width of the energy shell dE in the definition of the entropy does not change the classical entropy per particle S/N . Show that the choice of units in phase space does change the classical entropy per particle.

First consider a change in the width. We define the phase space shell volume as

$$\Omega(E)dE = \int_{E < \mathcal{H} < E+dE} d\mathbb{Q}d\mathbb{P}.$$

If we define

$$F(E) = \int_{\mathcal{H} < E} d\mathbb{Q}d\mathbb{P} = \int \Theta(E - \mathcal{H}) d\mathbb{Q}d\mathbb{P}$$

then we can write

$$\Omega(E)dE = F(E + dE) - F(E) \Rightarrow \Omega(E) = \frac{F(E + dE) - F(E)}{dE}.$$

In the limit that $dE \rightarrow 0$, this is the definition of a derivative and therefore is independent of dE .

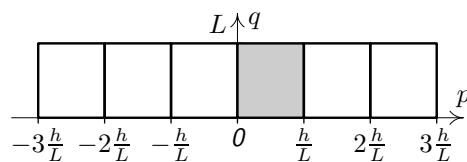
Regarding the units of phase space, if we take $\Omega \rightarrow \Omega/\sigma^{3N}$, then we get

$$\frac{S}{N} \rightarrow \frac{S}{N} - 3k_B \log \sigma.$$

Thus the entropy per particle shifts by a constant, independent of temperature.

(b) **Phase-space density of states for a particle in a one-dimensional box.** Show, or note, that the quantum momentum-space density of states for a free quantum particle in a one-dimensional box of length L with periodic boundary conditions is L/h . Draw a picture of the classical phase space of this box (p, x) , and draw a rectangle of length L for each quantum eigenstate. Is the phase-space area per eigenstate equal to h , as we assumed in 3.5?

A free particle has plane wave solutions $\psi(x) = e^{ikx}$, where k is determined by the boundary conditions. We have periodic conditions $\psi(0) = \psi(L)$ and $\psi'(0) = \psi'(L)$. Therefore we have $k_n = 2\pi n/L$. We have momentum $p = \hbar k = \hbar n/L$. Subtraction reveals a width $\Delta p = h/L$. More formally, we can differentiate to see that $g(p)dp \equiv dn = (L/h)dp$ so that the density of momentum states is $g(p) = L/h$. The quantum phase space is shown below.

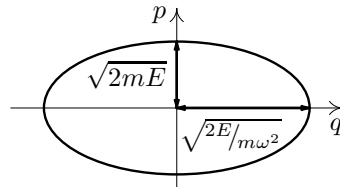


(c) **Phase-space density of states for N particles in a box.** Show that the density of states for N free particles in a cubical box of volume V with periodic boundary conditions is V^N/h^{3N} , and hence that the phase-space volume per state is h^{3N} .

The 3-D problem for N noninteracting particles is separable into $3N$ one-dimensional problems, so that as above, each particle has $\vec{p} = h\vec{n}/L$ where now \vec{n} has integer components. We find $\Delta\vec{p}$ is a cube with side-length h/L and therefore volume h^{3N}/V . This leads to a density $g(\vec{p}) = V/h^{3N}$ and the phase space volume is h^{3N} per state.

(d) **Phase-space density of states for a harmonic oscillator.** Consider a harmonic oscillator with Hamiltonian $\mathcal{H} = p^2/2m + \frac{1}{2}m\omega^2 q^2$. Draw a picture of the energy surface with energy E , and find the volume (area) of phase space enclosed. What is the volume per energy state, the volume between E_n and E_{n+1} , for the eigenenergies $E_n = (n + \frac{1}{2})\hbar\omega$?

Setting energy constant, we find an ellipse, pictured below:



The radii are found by setting $p = 0$ or $q = 0$, so that $p_{\max} = \sqrt{2mE}$ and $q_{\max} = \sqrt{2E/m\omega^2}$. The area of this ellipse is therefore

$$A = \pi p_{\max} q_{\max} = \frac{2\pi E}{\omega}.$$

For a quantum harmonic oscillator, $E_n = (n + \frac{1}{2})\hbar\omega$ so that

$$\Delta A = 2\pi\hbar = h.$$

So the phase space volume per state is again h .

7.4 Does Entropy Increase in Quantum Systems?.

Using the evolution law, prove that S is time-independent.

We have the density operator $\rho = \sum_i p_i |i\rangle\langle i|$, where $\mathcal{H}|i\rangle = E_i|i\rangle$ are energy eigenstates. Now we have from the text

$$\frac{d\rho}{dt} = \frac{1}{i\hbar} [\mathcal{H}, \rho] = \frac{1}{i\hbar} \sum_i (\mathcal{H}|i\rangle\langle i| - |i\rangle\langle i|\mathcal{H}) = \frac{1}{i\hbar} \sum_i (E_i|i\rangle\langle i| - |i\rangle\langle i|E_i) = 0,$$

since E_i are just numbers.

Since ρ is time-independent, it makes sense that $S = \text{Tr} \rho \log \rho$ is also time-independent, but we should be careful, since the different quantum pictures have very tricky time derivatives. We see that for a general operator \mathbf{A} ,

$$\frac{d}{dt} \text{Tr} \mathbf{A} = \sum_i \left[\left(\frac{d}{dt} \langle i| \right) \mathbf{A} |i\rangle + \langle i| \left(\frac{d\mathbf{A}}{dt} \right) |i\rangle + \langle i| \mathbf{A} \left(\frac{d}{dt} |i\rangle \right) \right].$$

We may choose to perform the trace in an energy eigenbasis, so that the derivatives of the state vectors are just $i\hbar \frac{d\langle i|}{dt} = -E_i \langle i|$ and $i\hbar \frac{d|i\rangle}{dt} = E_i |i\rangle$. Clearly these terms will cancel one another leaving $\frac{d}{dt} \text{Tr} \mathbf{A} = \text{Tr} \frac{d}{dt} \mathbf{A}$. Now we can verify that the entropy is constant,

$$\frac{d\boldsymbol{\rho}}{dt} = 0 \quad \Rightarrow \quad \frac{d}{dt} \boldsymbol{\rho} \log \boldsymbol{\rho} = 0 \quad \Rightarrow \quad \frac{dS}{dt} = 0.$$

Alternatively, we have:

$$\frac{dS}{dt} = \text{Tr} \left(\frac{d}{dt} \boldsymbol{\rho} \log \boldsymbol{\rho} \right) = \text{Tr} \left(\frac{d\boldsymbol{\rho}}{dt} \log \boldsymbol{\rho} + \frac{d\boldsymbol{\rho}}{dt} \right) = \text{Tr} \left(\frac{d\boldsymbol{\rho}}{dt} f(\boldsymbol{\rho}) \right),$$

where $f(\boldsymbol{\rho}) = \log \boldsymbol{\rho} + 1$. We know that density matrices in the Schrödinger picture evolve as $i\hbar \frac{d\boldsymbol{\rho}}{dt} = [\mathcal{H}, \boldsymbol{\rho}]$, so

$$i\hbar \frac{dS}{dt} = \text{Tr} [\mathcal{H} \boldsymbol{\rho} f(\boldsymbol{\rho})] - \text{Tr} [\boldsymbol{\rho} \mathcal{H} f(\boldsymbol{\rho})] = \text{Tr} [\mathcal{H} f(\boldsymbol{\rho}) \boldsymbol{\rho}] - \text{Tr} [\mathcal{H} f(\boldsymbol{\rho}) \boldsymbol{\rho}] = 0.$$

In the last step, we commute $\boldsymbol{\rho}$ with $f(\boldsymbol{\rho})$ in the first term and we cyclically permute the trace in the second.

Another, much simpler proof would be to observe that if we write everything in the Heisenberg picture, then all the terms in the trace are automatically time-independent, since only regular operators evolve. The density operator is made up of constant states in the Heisenberg picture and therefore everything in this question is constant.

7.5 Photon Density Matrices.

Write the density matrix for a photon linearly traveling along z and linearly polarized along \hat{x} , in the basis where $(1, 0)$ and $(0, 1)$ are polarized along \hat{x} and \hat{y} . Write the density matrix for a right-handed polarized photon, $(1/\sqrt{2}, i/\sqrt{2})$, and the density matrix for unpolarized light. Calculate $\text{Tr}(\boldsymbol{\rho})$, $\text{Tr}(\boldsymbol{\rho}^2)$, and $S = -k_B \text{Tr}(\boldsymbol{\rho} \log \boldsymbol{\rho})$. Interpret the values of the three traces physically: one is a check for pure states, one is a measure of information, and one is a normalization.

For a pure state with \hat{x} polarized photons we have

$$\boldsymbol{\rho}_x = |x\rangle\langle x| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

This gives $\text{Tr} \boldsymbol{\rho}_x = 1$. Squaring we find $\boldsymbol{\rho}_x^2 = \boldsymbol{\rho}_x$ so $\text{Tr} \boldsymbol{\rho}_x^2 = 1$ as well. Finally, $S = \text{Tr} \boldsymbol{\rho}_x \log \boldsymbol{\rho}_x = 0$ since $x \log x \rightarrow 0$ as $x \rightarrow 0$.

Now we consider a right-handed polarized state.

$$\boldsymbol{\rho}_R = |R\rangle\langle R| = \frac{1}{2}(|x\rangle + i|y\rangle)(\langle x| - i\langle y|) = \frac{1}{2} \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}.$$

Then $\text{Tr} \boldsymbol{\rho}_R = 1$, $\boldsymbol{\rho}_R^2 = \boldsymbol{\rho}_R$ so $\text{Tr} \boldsymbol{\rho}_R^2 = 1$. The entropy is a bit tricky, since we must diagonalize the matrix to take the log. Let

$$\mathbf{U} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}.$$

Then we see

$$\boldsymbol{\rho}_R = \mathbf{U} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \mathbf{U}^\dagger.$$

Traces are invariant under unitary transformations, so we see that this state has the same entropy as $\boldsymbol{\rho}_x$: $S = 0$. Finally we have the unpolarized state,

$$\boldsymbol{\rho}_u = \frac{1}{2}|x\rangle\langle x| + \frac{1}{2}|y\rangle\langle y| = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

This again has $\text{Tr} \rho_u = 1$, but now $\rho_u^2 = \frac{1}{2} \rho_u$ so $\text{Tr} \rho_u^2 = \frac{1}{2}$. This is not a pure state. The log is straightforward since the matrix is already diagonal and we see $S = \text{Tr} \rho_u \log \rho_u = k_B \log 2$.

So we have that $\text{Tr} \rho$ is a normalization, $\text{Tr} \rho^2$ is a check for pure states, and $S = \text{Tr} \rho \log \rho$ is a measure of information.

7.6 Spin Density Matrix.

Let the Hamiltonian for a spin be

$$\mathcal{H} = -\frac{\hbar}{2} \vec{B} \cdot \vec{\sigma}$$

where $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the three Pauli spin matrices, and \vec{B} may be interpreted as a magnetic field, in units where the gyromagnetic ratio is unity. Remember that $\sigma_i \sigma_j - \sigma_j \sigma_i = 2i\epsilon_{ijk} \sigma_k$. Show that any 2×2 density matrix may be written in the form

$$\rho = \frac{1}{2}(\mathbf{1} + \vec{p} \cdot \vec{\sigma}).$$

Show that the equations of motion for the density matrix $i\hbar \partial \rho / \partial t = [\mathcal{H}, \rho]$ can be written as $d\vec{p}/dt = -\vec{B} \times \vec{p}$.

Any 2x2 matrix can be written as

$$A = \frac{1}{2}(p_0 \mathbf{1} + \vec{p} \cdot \vec{\sigma}) = \frac{1}{2} \begin{pmatrix} p_0 + p_z & p_x - ip_y \\ p_x + ip_y & p_0 - p_z \end{pmatrix}.$$

It is easily verified that these four terms are independent. Moreover, the matrix is Hermitian if and only if p_0 and \vec{p} are real. Now, we know the density matrix evolves

$$\frac{d\rho}{dt} = \frac{1}{i\hbar} [\mathcal{H}, \rho].$$

Substituting $\mathcal{H} = -\frac{\hbar}{2} B_i \sigma_i$ and $\rho = \frac{1}{2}(\mathbf{1} + p_j \sigma_j)$ we see that the $\mathbf{1}$ commutes, leaving

$$\frac{d\rho}{dt} = -\frac{1}{i\hbar} \frac{\hbar}{4} B_i p_j [\sigma_i, \sigma_j] = -\frac{1}{4i} B_i p_j 2i\epsilon_{ijk} \sigma_k = -\frac{1}{2} (\vec{B} \times \vec{p}) \cdot \vec{\sigma}.$$

We can differentiate our original expression for ρ as well:

$$\frac{d\rho}{dt} = \frac{d}{dt} \frac{1}{2} (\mathbf{1} + \vec{p} \cdot \vec{\sigma}) = \frac{1}{2} \left(\frac{d\vec{p}}{dt} \right) \cdot \vec{\sigma}.$$

Since $\{\sigma_i\}$ are independent, we can conclude that

$$\frac{d\vec{p}}{dt} = -\vec{B} \times \vec{p}.$$

7.8 Einstein's A and B.

(a) Compute the long-time average ratio N_2/N_1 assuming only absorption and spontaneous emission. Even in the limit of weak coupling (small A and B), show that this equation is incompatible with the statistical distributions [7.83] and [7.84].

We have two states with energies $E_1 < E_2$ and populations N_1 and N_2 . We consider transition rates

$$\Gamma_{\text{abs}} = \Gamma_{1 \rightarrow 2} = 2\pi B u(\omega)$$

$$\Gamma_{\text{spon}} = \Gamma_{2 \rightarrow 1} = A$$

for photon absorption and spontaneous emission respectively.

The populations satisfy this differential equation:

$$\begin{pmatrix} \dot{N}_1 \\ \dot{N}_2 \end{pmatrix} = \frac{d}{dt} \begin{pmatrix} N_1 \\ N_2 \end{pmatrix} = \begin{pmatrix} -\Gamma_{1 \rightarrow 2} & \Gamma_{2 \rightarrow 1} \\ \Gamma_{1 \rightarrow 2} & -\Gamma_{2 \rightarrow 1} \end{pmatrix} \begin{pmatrix} N_1 \\ N_2 \end{pmatrix}.$$

The steady state solution $\dot{N}_1 = \dot{N}_2 = 0$ is

$$-\Gamma_{\text{abs}} N_1 + \Gamma_{\text{spon}} N_2 = 0 \quad \Rightarrow \quad 2\pi B u(\omega) N_1 = A N_2.$$

This gives an equilibrium ratio

$$\left\langle \frac{N_2}{N_1} \right\rangle = \frac{2\pi B}{A} u(\omega) = \frac{2\pi B}{A} \frac{\hbar}{\pi^2 c^3} \frac{\omega^3}{e^{\beta \hbar \omega} - 1}.$$

Einstein noted that B and A were independent of temperature so the only temperature dependence in this equation is the last denominator. We compare this to a Boltzmann distribution,

$$\left\langle \frac{N_2}{N_1} \right\rangle_{\text{Boltzmann}} = e^{-\beta(E_2 - E_1)} = e^{-\beta \hbar \omega}.$$

Setting these ratios equal, we can group all the temperature dependence on one side,

$$e^{-\beta \hbar \omega} (e^{\beta \hbar \omega} - 1) = 1 - e^{-\beta \hbar \omega} \stackrel{?}{=} \frac{2\pi B}{A} \frac{\hbar \omega^3}{\pi^2 c^3}.$$

This cannot generally be true. Therefore we must consider stimulated emission.

(b) Write the equation for dN_1/dt , including absorption (a negative term) and spontaneous and stimulated emission from the population N_2 . Assuming equilibrium, use this equation and equations 7.83 and 7.84 to solve for B , and B' in terms of A .

Now we add a rate $\Gamma_{\text{stim}} = 2\pi B' u(\omega)$ to $\Gamma_{2 \rightarrow 1}$, so that our new differential equation for \dot{N}_1 is

$$\frac{dN_1}{dt} = -2\pi B u(\omega) N_1 + (A + 2\pi B' u(\omega)) N_2 = N_2 \left(-\frac{2\pi \hbar \omega^3}{\pi^2 c^3} \cdot \frac{-B e^{\beta \hbar \omega} + B'}{e^{\beta \hbar \omega} - 1} + A \right) = 0,$$

where we now enforce the Boltzmann distribution by substituting $N_1 = N_2 e^{\beta \hbar \epsilon}$. We divide out N_2 and multiply through by $e^{\beta \hbar \omega} - 1$. Rearranging,

$$B' - B e^{\beta \hbar \omega} + A \left(\frac{\pi c^3}{2 \hbar \omega^3} \right) (e^{\beta \hbar \omega} - 1) = \left(\frac{\pi c^3}{2 \pi \hbar \omega^3} A - B \right) e^{\beta \hbar \omega} + \left(B' - \frac{\pi c^3}{2 \hbar \omega^3} A \right) = 0.$$

We group all the temperature dependence into the one exponential so that both parenthetical expressions must vanish. Thus,

$$B = \frac{\pi c^3}{2 \hbar \omega^3} A \quad B' = \frac{\pi c^3}{2 \hbar \omega^3} A.$$

(c) Show that the rate of decay of excited atoms $A + 2\pi B' u(\omega)$ is enhanced by a factor of $\langle n \rangle + 1$ over the zero temperature rate, where $\langle n \rangle$ is the expected number of photons in a mode at frequency $\hbar \omega = E_2 - E_1$.

We have $\Gamma_{\text{emit}} = \Gamma_{\text{spon}} + \Gamma_{\text{stim}} = A + 2\pi B' u(\omega)$. At $T = 0$, $u(\omega) \rightarrow 0$ so that $\Gamma_{\text{emit}}|_{T=0} = A$. Thus,

$$\frac{\Gamma_{\text{emit}}|_T}{\Gamma_{\text{emit}}|_{T=0}} = 1 + \frac{2\pi^2 c^3}{2 \hbar \omega^3} \frac{\hbar \omega^3}{\pi^2 c^3} \frac{1}{e^{\beta \hbar \omega} - 1} = 1 + \frac{1}{e^{\beta \hbar \omega} - 1} = 1 + \langle n \rangle,$$

where we recognize $\langle n \rangle = \frac{1}{e^{\beta \hbar \omega} - 1}$ as the expected number of photons (bosons) in the $\hbar \omega$ mode. Thus the total decay rate is enhanced by $n + 1$ over the zero-temperature rate.

7.9 Bosons are Gregarious: Superfluids and Lasers.

(a) **Adding a particle to a Bose condensate.** Calculate the symmetrized initial state of the system with the injected particle. Show that the ratio of the probability that the new boson enters the ground state (ψ_0) is enhanced over that of its entering an empty state (ψ_ℓ for $0 < \ell < M$) by a factor of $N + 1$.

This is a very tricky problem primarily because it is not particularly obvious how to properly symmetrize the final state $\Phi(r_1, \dots, r_{N+1}) = \psi_0(r_1) \dots \psi_0(r_N) \phi(r_{N+1})$ since $\phi = \frac{1}{\sqrt{M}} \sum_\ell \psi_\ell$ is not orthogonal to ψ_0 . Let's define the symmetrization operator S which is an honest linear operator,

$$S\psi(r_1, \dots, r_N) \equiv \frac{1}{N!} \sum_{\sigma \in S_N} \psi(r_{\sigma(1)}, \dots, r_{\sigma(N)}),$$

where S_N is the set of permutations. Extending the notation given in the problem, we define

$$|\Psi_{N+1}^{[0]}\rangle = |\psi_0(r_1) \dots \psi_0(r_{N+1})\rangle$$

$$|\Psi_{N+1}^{[\ell]}\rangle = \frac{1}{\sqrt{N+1}} \sum_{n=1}^{N+1} |\psi_0(r_1) \dots \psi_\ell(r_n) \dots \psi_0(r_{N+1})\rangle \quad (\ell \neq 0)$$

It is easy to see that $S|\Psi_{N+1}^{[\ell]}\rangle = |\Psi_{N+1}^{[\ell]}\rangle$ and that these states are orthonormal. We can now define the state $|\Phi\rangle$:

$$|\Phi\rangle = cS[|\Psi_N^{[0]}\rangle|\phi\rangle]$$

where c is a normalization constant. Using the definition of S , we obtain:

$$|\Phi\rangle = \frac{c}{\sqrt{M}} \left[|\Psi_{N+1}^{[0]}\rangle + \sum_{\ell=1}^{M-1} \frac{1}{\sqrt{N+1}} |\Psi_{N+1}^{[\ell]}\rangle \right]$$

Therefore the probability ratio of $\left| \langle \Psi_{N+1}^{[0]} | \Phi \rangle \right|^2$ to $\left| \langle \Psi_{N+1}^{[\ell]} | \Phi \rangle \right|^2$ is precisely $N+1$ for each $\ell \neq 0$.

(b) **Adding a photon to a laser beam.** Assuming a large box and a small decay rate Γ , find a formula for the number of modes M per unit volume V competing for the photon emitted from our atom. Evaluate your formula for a laser with wavelength $\lambda = 619\text{nm}$ and the line-width $\Gamma = 10^4\text{rad/s}$.

Consider a box with periodic boundary conditions and volume $V = L^3$. According to the problem statement, a decay rate Γ gives us an energy uncertainty $\Delta E \propto \hbar\Gamma$ and therefore an equal probability of decaying into any mode between $E - \hbar\Gamma/2 < \hbar\omega < E + \hbar\Gamma/2$. We need to know the density of states - the number of modes in this range. The wave number of quantized as $\vec{k} = \frac{2\pi}{L}\vec{n}$ where \vec{n} has integer components, and the energy is $\hbar\omega = \hbar ck$. The number of states with $\hbar\omega < E$ is given by

$$N(E) = 2 \cdot \frac{4}{3}\pi \left(\frac{L}{2\pi} k \right)^3 = \frac{V k^3}{3\pi^2} = \frac{V E^3}{3\pi^2 \hbar^3 c^3}$$

where the 2 is from counting both polarization states. Now

$$dN = g(E)dE = \frac{V E^2 dE}{\pi^2 \hbar^3 c^3}.$$

For small $\Delta E = \hbar\Gamma$, the number of modes in the shell, written in terms of $E = 2\pi\hbar c/\lambda$ is

$$M = g(E)\Delta E = \frac{V E^2 \Gamma}{\pi^2 \hbar^2 c^3} = \frac{4V\Gamma}{\lambda^2 c}.$$

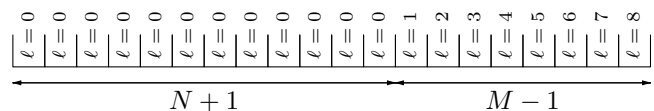
Substituting $\Gamma = 10^4\text{s}^{-1}$ and $\lambda = 619\text{nm}$ gives

$$\frac{M}{V} = 3.48 \times 10^8 \text{m}^{-3}.$$

(c) Using your result from part (a), give a formula for the number of photons per unit volume N/V there must be in the lasing mode for the atom to have a 50% likelihood of emitting into that mode.

We have M modes in our cavity, one of which is lasing. If $M-1$ of these modes (call them $1, \dots, M-1$) are empty and one mode (call this 0) has occupation N , then from (a) we expect a probability $P(0) \propto N+1$ and $P(\ell \neq 0) \propto 1$. But since there are $M-1$ of these modes (and the probabilities are exclusive), $P(1 \dots M-1) \propto M-1$. These two probabilities are mutual (one or the other will happen) and so add to one: $P(0) + P(1 \dots M-1) = 1$. This gives us the normalization:

$$P(0) = \frac{N+1}{(N+1) + (M-1)}.$$



For this probability to be $\frac{1}{2}$, we need

$$N+1 = M-1 \quad \Rightarrow \quad \frac{N}{V} = \frac{M}{V} - \frac{2}{V}.$$

For large boxes, as $V \rightarrow \infty$ we get

$$\frac{N}{V} \approx \frac{M}{V} = \frac{4\Gamma}{\lambda^2 c}.$$

7.10 Crystal Defects.

Write the canonical partition function $Z(T)$, the mean energy $E(T)$, the fluctuations in the energy, the entropy $S(T)$, and the specific heat $C(T)$ as a function of temperature. Plot the specific heat per defect $C(T)/N$ for $M = 6$; set the unit of energy equal to ϵ and $k_B = 1$ for your plot. Derive a simple relation between M and the change in entropy between zero and infinite temperature. Check this relation using your formula for $S(T)$.

We have N defects, each of which has one ground state of zero energy and M excited states with energy ϵ . The one-particle partition function is then

$$Z_1 = 1 + Me^{-\beta\epsilon},$$

so that the N -particle partition function is

$$Z = Z_1^N = (1 + Me^{-\beta\epsilon})^N.$$

We differentiate this to find the energy

$$E = -\frac{\partial}{\partial\beta} \log Z = -\frac{1}{Z} \frac{\partial Z}{\partial\beta} = \frac{NM\epsilon e^{-\beta\epsilon}}{1 + Me^{-\beta\epsilon}} = \frac{NM\epsilon}{M + e^{\beta\epsilon}}.$$

The fluctuations are given by (Eqs. 6.12 and 13)

$$\sigma_E^2 = \langle E^2 \rangle - \langle E \rangle^2 = \frac{\partial^2}{\partial\beta^2} \log Z = -\frac{\partial \langle E \rangle}{\partial\beta} = \frac{MN\epsilon^2 e^{\beta\epsilon}}{(M + e^{\beta\epsilon})^2}$$

$$\sigma_E = \frac{\epsilon\sqrt{MN}e^{\frac{1}{2}\beta\epsilon}}{M + e^{\beta\epsilon}}.$$

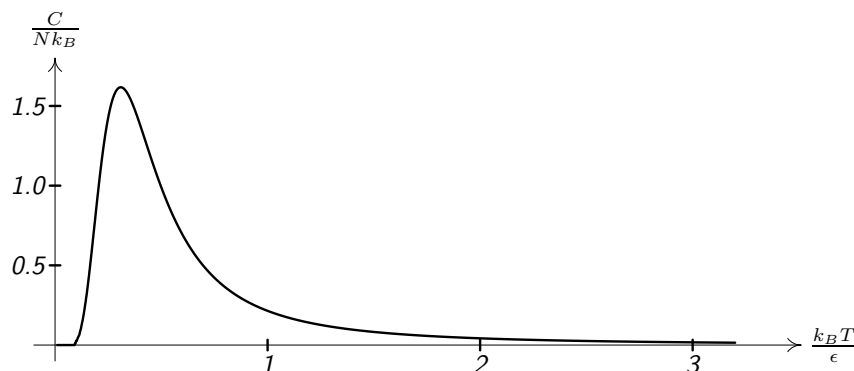
We find the entropy using $A = -k_B T \log Z$ and

$$S = -\frac{\partial A}{\partial T} = k_B (\log Z + \beta E) = Nk_B \log(1 + Me^{-\beta\epsilon}) + Nk_B \frac{NM\beta\epsilon}{M + e^{\beta\epsilon}}.$$

Finally, the specific heat is

$$C = \frac{\partial E}{\partial T} = -\frac{1}{k_B T^2} \frac{\partial E}{\partial\beta} = \frac{1}{k_B T^2} \sigma_E^2,$$

plotted below



Finally, we note that in the ground state there is no degeneracy in the energy levels so the entropy should vanish. At high temperatures, all the $M + 1$ states are equally likely so that the entropy should go to $Nk_B \log(M + 1)$. We confirm from our equation that

$$S(\infty) - S(0) = Nk_B \log(M + 1).$$

7.11 Phonons on a String.

Write the kinetic ...

Write the lagrangian as an integral over the lagrangian density, make the given change of variables to normal modes, and use the orthogonality of the sines and cosines to find:

$$\begin{aligned} L(t) &= \int_0^L dx \mathcal{L}(x, t) \\ &= \int_0^L dx \left[\frac{\mu}{2} \left(\sum_{m,n} \dot{q}_{k_n} \dot{q}_{k_m} \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{m\pi x}{L}\right) \right) \right. \\ &\quad \left. - \frac{\tau\pi^2}{2L^2} \left(\sum_{m,n} m n q_{k_n} q_{k_m} \cos\left(\frac{n\pi x}{L}\right) \cos\left(\frac{m\pi x}{L}\right) \right) \right] \\ &= \dots = \sum_{n \in \mathbb{Z}^+} \left[\frac{1}{2} \left(\frac{\mu L}{2} \right) \dot{q}_{k_n}^2 - \frac{1}{2} \left(\frac{\tau n^2 \pi^2}{2L} \right) q_{k_n}^2 \right]. \end{aligned}$$

This lagrangian describes a system of uncoupled harmonic oscillators with masses $\mu L/2$ and spring constants $n^2 \pi^2 \tau / 2L$. The oscillator frequencies are:

$$\omega_n = \omega_{k_n} = \frac{n\pi}{L} \sqrt{\frac{\tau}{\mu}} = \frac{n\pi c}{L},$$

where c is the wave speed. The density of states per unit frequency, as suggested in the problem, is the number of states in some frequency range ϵ that is large compared with the spacing between states $\Delta\omega = \pi c/L$, divided by ϵ . Since the spacing is uniform, the density of states is just $1/\Delta\omega$, or:

$$g(\omega) = \frac{1}{\Delta\omega} = \frac{L}{\pi c}.$$

Treating the oscillators quantum mechanically, the average occupancy of the n^{th} mode is:

$$\langle N_n \rangle = \frac{1}{e^{\beta \hbar \omega_n} - 1}.$$

The contribution of the n^{th} mode to the specific heat is, at fixed L (alternatively, one can find the total energy first and then differentiate to find the specific heat):

$$C_L^n = \frac{d\langle E_n \rangle}{dT} = \frac{d}{dT} \hbar \omega_n \left(\langle N_n \rangle + \frac{1}{2} \right),$$

and the specific heat per length is (at fixed L):

$$\begin{aligned} C_l &\equiv \frac{C_L}{L} = \frac{1}{L} \sum_n \frac{(\hbar \omega_n)^2}{k_B T^2} \frac{e^{\beta \hbar \omega_n}}{(e^{\beta \hbar \omega_n} - 1)^2} \\ &= \frac{\hbar^2}{\pi c k_B T^2} \sum_n \frac{\omega_n^2 e^{\beta \hbar \omega_n}}{(e^{\beta \hbar \omega_n} - 1)^2} \Delta\omega. \end{aligned}$$

In the limit as $L \rightarrow \infty$, $\Delta\omega \rightarrow 0$ and we can convert the sum to an integral:

$$\mathcal{C}_l \rightarrow \frac{\hbar^2}{\pi c k_B T^2} \int_0^\infty d\omega \frac{\omega^2 e^{\beta \hbar \omega}}{(e^{\beta \hbar \omega} - 1)^2}.$$

Using the fact (cf. integral table) that:

$$\int_0^\infty dx \frac{x}{e^x - 1} = \frac{\pi^2}{6},$$

we find:

$$\mathcal{C}_l = \frac{\pi k_B^2 T}{3 \hbar c}.$$

For a classical string, the equipartition theorem tells us that each harmonic or quadratic degree of freedom contributes $\frac{1}{2}k_B T$ to the internal energy. Each mode of the string therefore contributes k_B to the specific heat, so $\mathcal{C}_l \rightarrow \infty$, classically.

7.12 Semiconductors.

(a) Derive a formula for the number of electrons as a function of temperature T and chemical potential μ for the energy levels of our system.

For non-interacting fermions, the average occupation of a single-particle eigenstate is

$$\langle n_k \rangle = \frac{1}{e^{\beta(\epsilon_k - \mu)} + 1}.$$

The number of electrons in our system is fixed at $N + M$, so we can sum over $\langle n_k \rangle$ for the three energy levels,

$$N + M = \frac{N}{e^{-\beta(\Delta/2 + \mu)} + 1} + \frac{M}{e^{\beta(\Delta/2 - \epsilon - \mu)} + 1} + \frac{N - M}{e^{\beta(\Delta/2 - \mu)} + 1}$$

We multiplied the single-particle occupations by the number of states in each level. Now we can solve for $\mu(T)$.

(b) What is the limiting occupation probability for the states as $T \rightarrow \infty$, where entropy is maximized and all states are equally likely? Using this, find a formula for $\mu(T)$ valid at large T , not involving Δ or ϵ .

The limiting occupation probability as $T \rightarrow \infty$ is $p = (N + M)/2N$ since there are $N + M$ electrons distributed in this case with equal probability among $2N$ states. The average number of electrons in a single-particle state is the same, giving us a relation for μ :

$$\frac{N + M}{2N} = \frac{1}{e^{-\beta\mu} + 1},$$

where we have made the assumption that for large T , $e^{-\beta\epsilon_k} \approx 1$. Solving for μ one finds

$$\mu \approx k_B T \log \left(\frac{N + M}{N - M} \right).$$

(c) Draw an energy level diagram showing the filled and empty states at $T = 0$. Find a formula for $\mu(T)$ in the low temperature limit $T \rightarrow 0$, not involving the variable T .

At high temperatures, it was easy to solve for μ , but at low temperatures, we have a problem. As $T \rightarrow 0$, the total occupancy $N(\mu)$ approaches a step function, pictured below. We see that the problem arises because most of the interval $\Delta/2 - \epsilon < \mu < \Delta/2$ has total occupancy very close to $N + M$. We must find the value of μ that crosses $N_{\text{tot}} = N + M$ exactly. We ignore the valence band because the energy gap $\Delta \gg k_B T$ at low temperatures, so electrons are very unlikely to be excited across it. This leaves M electrons to distribute between the conduction and impurity bands,

$$\frac{M}{e^{\beta(\frac{\Delta}{2} - \epsilon - \mu)} - 1} + \frac{N - M}{e^{\beta(\frac{\Delta}{2} - \mu)} - 1} = M.$$

(Alternately we could equate the number of electrons in the conduction band with the number of holes (vacancies) in the impurity band.) We multiply both sides by the product of the denominators and solve a quadratic equation in $e^{-\beta\mu}$ to find

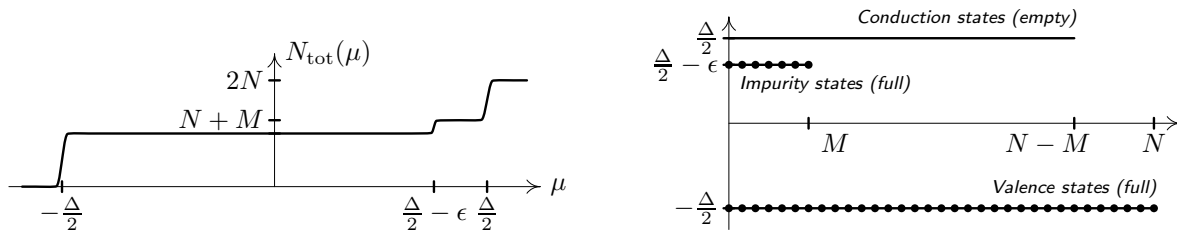
$$e^{\beta(\frac{\Delta}{2} - \mu)} = \frac{N - 2M}{2M} + \frac{\sqrt{(N - 2M)^2 + 4(N - M)Me^{\beta\epsilon}}}{2M}.$$

For large β , only the exponential will remain on the right-hand side, leaving

$$e^{-\beta\mu} = \left(2\sqrt{\frac{N}{M}} - 1\right) e^{-\beta(\frac{\Delta}{2} - \frac{\epsilon}{2})}.$$

When we take the log, the $2\sqrt{N/M} - 1$ term will vanish in light of the β (we may need very small T if $N \gg M$ - but we're doing theory so we can make it arbitrarily cold!), so that

$$\mu = \frac{\Delta}{2} - \frac{\epsilon}{2}.$$



(d) In a one centimeter cubed sample, there are $M = 10^{16}$ phosphorus atoms; silicon has about $N = 5 \times 10^{22}$ atoms per cubic centimeter. Find μ at room temperature (1/40eV) from the formula you derived in part (a). At this temperature, what fraction of the phosphorus atoms are ionized? What is the density of holes?

To find μ from the formula in part (a) with $T = T_{\text{room}}$, Mathematica's 'FindRoot' function gives $\mu \approx .194376 \text{ eV} \approx .2 \text{ eV}$. The fraction of phosphorous atoms ionized is the fraction of states at energy $\Delta/2 - \epsilon$ that are unoccupied:

$$f = 1 - \frac{1}{e^{\beta(\Delta/2 - \epsilon - \mu)} + 1} = \dots \approx .9999,$$

so essentially all of the them are ionized at room temperature. The density of holes is

$$\begin{aligned} \rho_{\text{holes}} &= (\text{fraction of unoccupied states at } -\Delta/2) \cdot (\text{number density of levels at } -\Delta/2) \\ &= \left(1 - \frac{1}{e^{-\beta(\Delta/2 + \mu)} + 1}\right) \cdot (N) \approx 1.77 \times 10^9 \text{ cm}^{-3} \approx 2 \times 10^9 \text{ cm}^{-3}. \end{aligned}$$

7.13 Bose Condensation in a Band.

Which of the following is necessary to undergo Bose condensation at low temperatures?

- (a) $g(E)/(e^{\beta(E-E_{\min})} + 1)$ is finite as $E \rightarrow E_{\min}^-$.
- (b) $g(E)/(e^{\beta(E-E_{\min})} - 1)$ is finite as $E \rightarrow E_{\min}^-$.
- (c) $E_{\min} \geq 0$.
- (d) $\int_{E_{\min}}^E g(E')/(E' - E_{\min})dE'$ is convergent at the lower limit E_{\min} .
- (e) Bose condensation cannot occur in a system whose states are confined to an energy band.

The number of particles in the system is (with a continuum approximation for $g(E)$)

$$N = \int_{E_{\min}}^{E_{\max}} dE \frac{g(E)}{e^{\beta(E-\mu)} - 1}.$$

Now, $\langle n_E \rangle \rightarrow \infty$ as $\mu \rightarrow E^-$ and $\langle n_E \rangle$ becomes negative when $\mu > E$, so it must be true that $\mu \leq E_{\min}$. This sets an upper bound on the number of particles in the 'continuum-approximated' system—i.e. we get N_{\max} when $\mu \rightarrow E_{\min}^-$. As long as particle number is conserved, particles in excess of N_{\max} must go into the ground state (or close to it), where the continuum approximation fails. Thus, the necessary condition for having Bose condensation at low temperatures is that

$$N_{\max} = \int_{E_{\min}}^{E_{\max}} dE \frac{g(E)}{e^{\beta(E-E_{\min})} - 1} < \infty.$$

This integral is only in danger of diverging at the lower limit, so we will look only at the integral in that neighborhood. The integral converges if and only if, for any $\delta > 0$ (arbitrarily small),

$$\int_{E_{\min}}^{E_{\min}+\delta} dE \frac{g(E)}{e^{\beta(E-E_{\min})} - 1} < \infty.$$

For small enough δ , the denominator in the integrand is

$$e^{\beta(E-E_{\min})} - 1 \approx \beta(E - E_{\min}).$$

Thus, a necessary condition for the system to undergo Bose condensation is

$$\int_{E_{\min}}^E dE \frac{g(E)}{E - E_{\min}} < \infty \Rightarrow \boxed{\text{D.}}$$

The others are not generally true.

7.15 The Photon-dominated Universe.

- (a) What is w' ? What is $n'(w')$? Show that Plank's form is unchanged; what is the new temperature T' ?

$$\begin{aligned} w' &= \frac{w}{f} \\ n'(w') &= \frac{n(w)}{f^3} \\ n'(w')dw' &= \frac{n(w)}{f^3}dw \rightarrow n'(w') = n(w', T') \\ T' &= \frac{T}{f} \end{aligned}$$

(b) How many microwave background photons are there per cubic centimeter?

$$n_{ph} = \frac{1}{\pi^2 c^3} \left(\frac{k_B T}{\hbar} \right)^3 \int dx \frac{x^2}{e^x - 1}$$

$$n_{ph} = 4.14 \times 10^2 \text{ atoms/cm}^3 \gg n_{matter}$$

(c) Calculate S,E,P for the photon gas.

$$S = \frac{16}{3} \frac{V \sigma T^3}{c}$$

$$E = 4 \frac{V \sigma T^4}{c}$$

$$P = \frac{4}{3} \frac{\sigma T^4}{c}$$

(d) Calculate S,E,P for an ideal gas of hydrogen atoms.

We can ignore quantum mechanics when treating this gas because its density is much smaller than the quantum density. We have:

$$A = Nmc^2 + Nk_B T (\ln(n\lambda^3) - 1)$$

$$S = Nk_B \left(-\ln(n\lambda^3) + \frac{5}{2} \right)$$

$$E = Nmc^2 + \frac{3}{2} Nk_B T$$

$$P = nk_B T$$

Plugging in the numbers, we have:

	<i>matter</i>	<i>photons</i>
$\frac{S}{V} \left(\frac{\text{erg}}{K \text{ cm}^3} \right)$	2.27×10^{-21}	8.2×10^{-11}
$\frac{E}{V} \left(\frac{\text{erg}}{\text{cm}^3} \right)$	3.8×10^{-10}	1.7×10^{-10}
$P \left(\frac{\text{erg}}{\text{cm}^3} \right)$	9.3×10^{-23}	1.2×10^{-12}

Hence, matter dominates the internal energy but

photons dominate the rest.

(e) Show that the sum of the two photon contributions to the mass density is proportional to $\frac{E}{c^2 V}$. What is the constant of proportionality?

$$\rho_{ph} = \frac{4}{3} \frac{E}{c^2 V}$$

(f) Find B.

$$B = \frac{4}{9} \frac{E}{V}$$

(g) Find the speed of sound.

$$v = \frac{c}{\sqrt{3(R+1)}}$$

7.16 White Dwarves, Neutron Stars, and Black Holes.

(a) Assuming non-relativistic electrons, calculate the energy of a sphere with N zero-temperature non-interacting electrons and radius R . Calculate the Newtonian gravitational energy of a sphere of He^4 nuclei of equal and opposite charge density. At what radius is the total energy minimized?

For a free fermi gas of electrons, the density of states is (cf. eq. 7.74 in draft text) $g(\epsilon) = \frac{\sqrt{2} V m_e^{3/2}}{\pi^2 \hbar^2} \sqrt{\epsilon}$. At zero temperature, the fermi energy is found from the constraint on the number of particles $N_e = \int_0^{\epsilon_F} g(\epsilon) d\epsilon$. Since the star is electrically neutral, we can approximate $N_e \approx M/2m_p$. Solving for ϵ_F we then calculate

$$\epsilon_F = \frac{1}{8} \left(\frac{9M\pi}{m_p} \right)^{2/3} \frac{\hbar^2}{m_e R^2}.$$

Now we can compute the energy of the electron gas at $T = 0$

$$E_e = \int_0^{\epsilon_F} \epsilon g(\epsilon) d\epsilon = \dots = \frac{1}{120\pi} \cdot \frac{\hbar^2}{m_e R^2} \left(\frac{9M\pi}{m_p} \right)^{5/3}.$$

One way to find the Newtonian gravitational energy of a sphere is to imagine building it up from the center out by adding small shells of mass. We integrate over the potential energy of each 'shell' of mass interacting with the mass inside:

$$E_g = -G\rho^2 \int_0^R dr (4\pi r^2) \left(\frac{4\pi}{3} r^3 \right) \frac{1}{r} = -\frac{3GM^2}{5R}.$$

The radius at which the energy is extremized is found in the usual way, with $E_{\text{tot}} = E_e + E_g$:

$$\frac{\partial E_{\text{tot}}}{\partial R} = 0 \iff R_{\text{min}} = \frac{(9\pi)^{2/3} \hbar^2}{8Gm_e m_p^{5/3} M^{1/3}}.$$

That it is a minimum is clear from a sketch of the energy, which goes as $1/R^2 - 1/R$.

(b) Using the non-relativistic model in part (a), calculate the Fermi energy of the electrons in a white dwarf star of the mass of the Sun, $2 \times 10^{33}\text{g}$, assuming that it is composed of helium. (i) Compare it to a typical chemical binding energy of an atom. Are we justified in ignoring the electron-electron and electron-nuclear interactions (i.e., chemistry)? (ii) Compare it to the temperature inside the star, say 10^7K . Are we justified in assuming that the electron gas is degenerate (roughly zero temperature)? (iii) Compare it to the mass of the electron. Are we justified in using a non-relativistic theory? (iv) Compare it to the mass difference between a proton and a neutron.

Putting in numbers for a star of mass $2 \times 10^{33}\text{g}$ we find $R_{\text{min}} \approx 7150\text{km}$ so that

$$\epsilon_F(R_{\text{min}}) = \dots \approx 190\text{keV}.$$

(i) Typical chemical binding energies in atoms are about $1 - 100\text{eV} \ll \epsilon_F$, so we are justified in treating the nuclei and electrons as non-interacting.

(ii) At a temperature of 10^7K , $k_B T \approx 1\text{keV} \ll \epsilon_F$, justifying the zero temperature approximation.

(iii) $m_e c^2 \approx 511\text{keV} \approx 3\epsilon_F$, so relativistic effects shouldn't be ignored.

(iv) $(m_p - m_n)c^2 \approx 1.3\text{MeV} \approx 7\epsilon_F$, so the process of electrons and protons combining to form neutrons should be largely suppressed, but could be important.

(c) Assuming extremely relativistic electrons with $\epsilon = pc$, calculate the energy of a sphere of non-interacting electrons. Notice that this energy cannot balance against the gravitational energy of the nuclei except for a special value of the mass, M_0 . Calculate M_0 . How does your M_0 compare with the mass of the Sun, above?

Assuming extremely relativistic electrons with $\epsilon = pc$, the density of states is

$$g(\epsilon) = \frac{V}{\pi^2 (\hbar c)^3} \epsilon^2.$$

The same calculations as above now yield

$$\epsilon_F = \left(\frac{3M\pi^2 (\hbar c)^3}{2m_p V} \right)^{1/3}$$

$$\text{and } E_e = \frac{3^{5/3} \hbar c \pi^{1/3}}{2^4 R} \left(\frac{M}{m_p} \right)^{4/3}.$$

This time E_e and E_g both go as $1/R$, so there's no extremum except when they are equal and opposite, in which case the energy $E_e + E_g$ is identically zero. This happens when the mass is:

$$M_0 = \frac{3\pi^{1/2}}{2^6 m_p^2} \left(\frac{5\hbar c}{G} \right)^{3/2} \approx 3.44 \times 10^{30}\text{kg} \approx 1.7M_\odot.$$

8.2 Ising Fluctuations and Susceptibilities.

(a) Show that the average of the magnetization M equals $-(\partial\mathcal{F}/\partial H)|_T$. Derive the formula writing the susceptibility $\chi_0 = (\partial M/\partial H)|_T$ in terms of $\langle(M - \langle M \rangle)^2\rangle = \langle M^2 \rangle - \langle M \rangle^2$.

First note that the magnetization of the n^{th} state is

$$M_n = \sum_i s_i = -\frac{\partial E_n}{\partial H}.$$

Differentiating the free energy gives us

$$-\frac{\partial\mathcal{F}}{\partial H}\bigg|_T = \frac{k_B T}{Z} \sum_n e^{-\beta E_n} \left(-\beta \frac{\partial E_n}{\partial H}\right) = \frac{1}{Z} \sum_n e^{-\beta E_n} M_n = \langle M \rangle.$$

Finding the formula for the susceptibility is straightforward:

$$\chi_0 = \frac{\partial \langle M \rangle}{\partial H}\bigg|_T = -\frac{\partial^2 \mathcal{F}}{\partial H^2}\bigg|_T = \dots = \beta \left(\langle M^2 \rangle - \langle M \rangle^2 \right).$$

(b) **Correlations and Susceptibilities: Numerical.** Check the formulas for C and χ from part (b) at $H = 0$ and $T = 3$ by measuring the fluctuations and the averages, and then changing by $\Delta H = 0.02$ or $\Delta T = 0.1$ and measuring the averages again. Check them also for $T = 2$, where $\langle M \rangle \neq 0$.

Careful measurement should yield results that agree to within 1%.

(c) **Low Temperature Expansion for the Magnetization.** What is the energy for flipping a spin antiparallel to its neighbors? Equilibrate at low temperature $T = 1.0$, and measure the magnetization. Notice that the primary excitations are single spin flips. In the low temperature approximation that the flipped spins are dilute, write a formula for the magnetization. Check your prediction against the simulation.

The energy for flipping a (initially parallel) spin antiparallel to its neighbors is

$$\Delta E = 8J + 2H,$$

assuming the spin was already aligned with H . In the dilute spin approximation valid at low temperatures, we can find the magnetization by first finding the magnetization $\langle M_1 \rangle$ for a single spin assuming all its neighbors are parallel to H . Then the total magnetization is simply

$$\langle M \rangle = N \langle M_1 \rangle = N \frac{e^{-\beta E_\uparrow} - e^{-\beta E_\downarrow}}{Z} = N \tanh(\beta(4J \pm H)),$$

which can be further approximated by keeping only first-order terms in $e^{-\beta\Delta E}$,

$$\langle M \rangle \approx N (1 - 2e^{-\beta\Delta E}).$$

This formula agrees well with the simulation at low temperatures.

(d) **High Temperature Expansion for the Susceptibility.** Calculate a formula for the susceptibility of a free spin coupled to an external field. Compare it to the susceptibility you measure at high temperature $T = 100$ for the Ising model (say, $\Delta M/\Delta H$ with $\Delta H = 1$. Why is $H = 1$ a small field in this case?)

For a free spin coupled to an external field,

$$\langle M \rangle = \frac{e^{-\beta E_{\uparrow}} - e^{-\beta E_{\downarrow}}}{Z} = \tanh(\beta H),$$

and

$$\chi = \left. \frac{\partial \langle M \rangle}{\partial H} \right|_T = \beta \operatorname{sech}^2(\beta H),$$

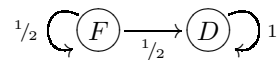
so that $\chi \rightarrow \beta$ as $T \rightarrow \infty$. This should agree well with the simulation at high temperatures.

8.3 Waiting for Godot, and Markov.

(The title of this exercise was a mistake. I was mixing up "Waiting for Godot", by Samuel Becket, with "Rosencrantz & Guildenstern Are Dead", by Tom Stoppard. It is the latter play that has the coin-flipping scene... JPS)

(a) Treat the two states of the physicist ("still flipping" and "done") as states in a Markov process. The current probability vector then is $\vec{\rho} = \begin{pmatrix} \rho_{\text{flipping}} \\ \rho_{\text{done}} \end{pmatrix}$. Write the transition matrix \mathcal{P} , giving the time evolution $\mathcal{P} \cdot \vec{\rho}_n = \vec{\rho}_{n+1}$, assuming that the coin is fair.

We can represent this processes with a diagram.



The transition matrix is therefore

$$\mathcal{P} = \begin{pmatrix} 1/2 & 0 \\ 1/2 & 1 \end{pmatrix}.$$

As a check we note that the elements of \mathcal{P} satisfy $0 \leq \mathcal{P}_{ij} \leq 1$, and the columns sum to 1, as they must for any acceptable transition matrix.

(b) Find the eigenvalues and right eigenvectors of \mathcal{P} . Which eigenvector is the steady state ρ^* ? Call the other eigenvector $\tilde{\rho}$. For convenience, normalize $\tilde{\rho}$ so that its first component equals one.

The eigenvalues and eigenvectors are

$$\lambda^* = 1, \quad \rho^* = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\tilde{\lambda} = 1/2, \quad \tilde{\rho} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

(c) Assume an arbitrary initial state is written $\rho_0 = A\rho^* + B\tilde{\rho}$. What are the conditions on A and B needed to make ρ_0 a valid probability distribution? Write ρ_n as a function of A and B , ρ^* and $\tilde{\rho}$.

If the initial state is written

$$\rho_0 = A\rho^* + B\tilde{\rho} = \begin{pmatrix} B \\ A - B \end{pmatrix},$$

then in order for this to be a valid probability distribution, A and B must satisfy

$$\sum_i \rho_i = 1 \Rightarrow A = 1,$$

$$0 \leq \rho_i \leq 1 \Rightarrow 0 \leq B \leq 1.$$

And after n flips,

$$\rho_n = \rho^* + \frac{B}{2^n} \tilde{\rho}.$$

8.4 Red and Green Bacteria.

(a) After a very long time, what is the probability distribution for the number α of red bacteria in the growth medium?

This is a Markov process, since the evolution of the system depends only on its current state and not on its history. It is not ergodic, however, because there are two states that cannot be left once they are reached – all green or all red – and given enough time, one of these states will inevitably be reached. Since the initial condition has an equal number of red and green bacteria, and since the predator is color-blind, the system will arrive at one or the other with equal probability. The state of the system can be specified by a single number, for instance the number of reds, which is a number from 0 to 1000. So we can write the state as a 1001-component vector with all 0s and one 1 in the position corresponding to the number of reds. We could use the same kind of vector to represent the probability distribution of an ensemble of systems. If we do this then the initial distribution is

$$\rho_0 = (\underbrace{0, \dots, 0}_{500}, 1, \underbrace{0, \dots, 0}_{500}).$$

The limiting distribution is then

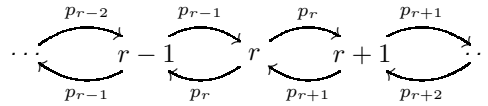
$$\rho(t \rightarrow \infty) = (1/2, 0, \dots, 0, 1/2).$$

(b) Roughly how long will it take to reach this final state?

There are a number of ways to approach this problem, but no exact analytical method that I know of (apart from the fact that the transition matrix is finite and could therefore be multiplied out by brute force). There are however a number of analytical approximation schemes that work very well, and writing a computer simulation to solve the problem is also a popular method.

Considering the case in which all the bacteria multiply and then 1000 are consumed, the evolution of the state is like a variable step size random walk. The step size in this case depends on the number of reds, for example, and is largest in the center and smaller as the system nears one of the equilibrium states (equivalently this is a variable-rate diffusion problem with interesting boundary conditions). The time to reach the end can be estimated by summing all the steps out from the center until they add up to 500.

Another approach is to consider the case in which 1 bacterium is eaten at a time. Let N be the total number of bacteria and r be the number of reds. Then the transition diagram is



so that the components of the transition matrix are

$$\mathcal{P}_{r+1 \leftarrow r} = \frac{r}{N} \frac{N-r}{N+1} \equiv p_r \quad (\text{red divides, green eaten})$$

$$\mathcal{P}_{r \leftarrow r} = \frac{r}{N} \frac{r+1}{N+1} + \frac{N-r}{N} \frac{N+1-r}{N+1} \quad (\text{same divides and is eaten})$$

$$\mathcal{P}_{r-1 \leftarrow r} = \frac{N-r}{N} \frac{r}{N+1} = \mathcal{P}_{r+1 \leftarrow r} \quad (\text{green divides, red eaten})$$

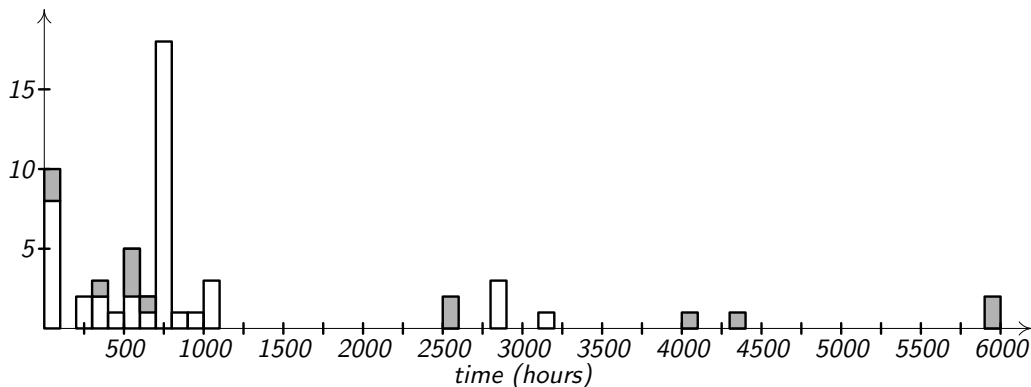
Note that the probabilities sum to one, $\sum_s \mathcal{P}_{sr} = 1$, as they must. One could now write down a difference equation for ρ_{n+1} and ρ_n and convert it to the diffusion equation, as some people did. Alternatively note that all the rows of the transition matrix add up to the same value, $\lambda = 1 - 2/N(N+1)$, except for the first and last rows. But since $\mathcal{P}_{0j} = \delta_{0j}$, λ is an eigenvalue of the matrix with the eigenvector

$$\rho_\lambda = (a, 1, 1, \dots, 1, 1, a),$$

where a is a constant determined from the first and last row condition. I don't know whether it can be proved, but we can surmise that this is the largest eigenvalue smaller than 1 and therefore the slowest decaying eigenmode. This is reasonable because the eigenvector is essentially flat, with no "nodes" (zeroes). This is what a diffusing system looks like after a long time, after the initial peaks and valleys have smoothed out. If our system were to begin in this eigenstate, then how long would it take to decay? After t hours the state would be:

$$\rho(t) \propto (\lambda)^{1000t} \rho_\lambda \approx e^{-\frac{2000t}{N(N+1)}} \rho_\lambda,$$

which decays with a half-life of $T_{1/2} \approx 347$ hours. Our system instead begins in a superposition of this state and many others and therefore the result will be different, but this is nonetheless a reasonable estimate. Solutions by computer simulation give a time of around 600 hours to reach the final state. Here's a histogram of the varying times that were reported in the past two years for your viewing enjoyment (this year is shaded):



Distribution of reported times for part (b).

(c) Assume that the predator has a 1% preference for green bacteria. Roughly how much will this change the final distribution?

In the diffusion equation or variable step size approach, a small preference for green bacteria adds an effective force that pushes the system towards the all red state. This has a very big impact on the final distribution, as can also be found computationally. The final distribution is around 99.9% red and < 1% green, and it reaches the final state in around 50-100 hours on average.

8.5 Detailed Balance.

(a) Assume that each of the three types of transitions among the three states satisfies detailed balance. Eliminate the equilibrium probability densities to write the unknown rate $P_{\alpha \leftarrow \beta}$ in terms of the five other rates.

We have a three-state system and assume detailed balance in the form

$$P_{\beta\alpha}\rho_\alpha^* = P_{\alpha\beta}\rho_\beta^*, \quad (1)$$

for each of the three types of transitions (i.e. $\alpha \leftrightarrow \beta$, $\alpha \leftrightarrow \gamma$, $\beta \leftrightarrow \gamma$). We can eliminate the three ρ_α s to find

$$P_{\alpha\beta}P_{\beta\gamma}P_{\gamma\alpha} = P_{\gamma\beta}P_{\beta\alpha}P_{\alpha\gamma}. \quad (2)$$

(b) Suppose P is the transition matrix for some Markov process satisfying the condition 8.28 for every triple of states α , β , and γ . Assume for simplicity that there is a state α with non-zero transition rates from all other states δ . Construct a probability density ρ_δ^* that demonstrates that P satisfies detailed balance.

Now we consider a system with an arbitrary number of states, such that any three of the states satisfy detailed balance in the form (2). We wish to construct a probability distribution ρ_δ^ that satisfies detailed balance in the form (1) between any two states, and we suppose that there is some specific state α that has non-zero transition rates from all other states. Guess:*

$$\rho_\delta^* = \frac{P_{\delta\alpha}}{P_{\alpha\delta}}\rho_\alpha^*.$$

This distribution satisfies detailed balance between the state α and any other state δ by construction. It is well-defined because the transition rates $P_{\alpha\delta}$ are non-zero by assumption. But does it satisfy (1) for any two states, say β and γ ? Yes, because

$$\rho_\beta^* = \frac{P_{\beta\alpha}}{P_{\alpha\beta}}\rho_\alpha^*, \quad \rho_\gamma^* = \frac{P_{\gamma\alpha}}{P_{\alpha\gamma}}\rho_\alpha^*,$$

and since (2) holds,

$$P_{\alpha\beta}P_{\beta\gamma}P_{\gamma\alpha} = P_{\alpha\gamma}P_{\gamma\beta}P_{\beta\alpha}.$$

Therefore,

$$P_{\beta\gamma}\rho_\gamma^* = P_{\gamma\beta}\rho_\beta^*.$$

8.6 Metropolis.

(a) Show that Metropolis satisfies detailed balance. Note that it is ergodic and Markovian (no memory), and hence that it will lead to thermal equilibrium. Is Metropolis more efficient than the heat-bath algorithm (fewer random numbers needed to get to equilibrium)?

The heat bath algorithm only flips one spin at a time, so $\mathcal{P}_{\alpha\beta} = 0$ unless the states α and β differ by at most one flipped spin. In that case,

$$\mathcal{P}_{\uparrow\leftrightarrow\downarrow}\rho_{\downarrow}^* = \frac{e^{-\beta E_{\uparrow}}}{e^{-\beta E_{\uparrow}} + e^{-\beta E_{\downarrow}}} \left(\frac{e^{-\beta E_{\downarrow}}}{Z} \right) = \mathcal{P}_{\downarrow\leftrightarrow\uparrow}\rho_{\uparrow}^*.$$

For the Metropolis algorithm, assume for the moment that $E_{\uparrow} > E_{\downarrow}$. Then $\Delta E > 0$ and

$$\mathcal{P}_{\uparrow\leftrightarrow\downarrow}\rho_{\downarrow}^* = e^{-\beta\Delta E} \cdot \frac{e^{-\beta E_{\downarrow}}}{Z} = \frac{e^{-\beta E_{\uparrow}}}{Z} = \mathcal{P}_{\downarrow\leftrightarrow\uparrow}\rho_{\uparrow}^*,$$

since for $\uparrow\Rightarrow\downarrow$, $\Delta E < 0$. The case with $E_{\downarrow} > E_{\uparrow}$ is analogous. In both of these algorithms any spin has a finite probability of being flipped, so any state of the system can be reached from any other state—i.e. they are ergodic (irreducible). Furthermore, they are Markovian since each change depends only on the current configuration of the system and not on its history. Therefore, the main theorem of chapter 8 (Theorem 8.4) applies and we conclude that both algorithms will reach thermal equilibrium.

The Metropolis algorithm is more efficient, as it requires on average fewer random numbers to reach equilibrium. One can see that its efficiency depends on the temperature as well. For example, at a low temperature (below T_c), the equilibrium state will be one with most of the spins parallel, a low energy state. Beginning from some arbitrary configuration of spins, most of the flips toward equilibrium will have $\Delta E < 0$. In this situation Metropolis would be considerably more efficient than the heat bath. At higher temperatures on the other hand, the efficiency gain is significantly reduced.

(b) Show for the two configurations shown above that $E_A - E_B = 2(n_{\uparrow} - n_{\downarrow})J$. Argue that this will be true for flipping any cluster of up spins to down.

8.8 Wolff.

(a) Show for the two configurations shown above that $E_A - E_B = 2(n_{\uparrow} - n_{\downarrow})J$. Argue that this will be true for flipping any cluster of up spins to down.

Assume $H = 0$. In flipping any cluster of up spins to down, the spin-spin interactions internal and external to the cluster remain unchanged. However, up spins outside the cluster and along the boundary will suffer a change of interaction energy $\Delta E = +2n_{\uparrow}J$. Here n_{\uparrow} is the total number of up spin ‘interactions’ on the boundary, so for example the cluster depicted in Fig. 8.10 in the text has $n_{\downarrow} = 6$ – one of the down spins gets counted twice since it’s interacting with two of the spins in the cluster. Similarly down spins outside the cluster along the boundary will contribute a change in energy $\Delta E = -2n_{\downarrow}J$. Therefore, in flipping any cluster of up spins to down, $E_A - E_B = 2(n_{\uparrow} - n_{\downarrow})J$.

(b) What value of p lets the Wolff algorithm satisfy detailed balance at temperature T ?

Detailed balance requires

$$\Gamma_{A\Rightarrow B}e^{-\beta E_A} = \Gamma_{B\Rightarrow A}e^{-\beta E_B}.$$

This is equivalent to

$$\sum_{\alpha} p_{\alpha}^c (1-p)^{n_{\uparrow}} = \sum_{\alpha} p_{\alpha}^c (1-p)^{n_{\downarrow}} e^{-\beta(E_B - E_A)}$$

$$\Rightarrow (1-p)^{n_{\uparrow} - n_{\downarrow}} = e^{-2\beta(n_{\uparrow} - n_{\downarrow})J}.$$

Therefore, provided that $n_{\uparrow} \neq n_{\downarrow}$, we conclude that

$$p = 1 - e^{-2\beta J}.$$

This value of p lets the Wolff algorithm satisfy detailed balance at temperature T including the specific case with $n_{\uparrow} = n_{\downarrow}$, although in that case any other value of p works too.

8.10 Stochastic Cells.

(a) **Continuum dimerization.** Write the differential equation for dM/dt treating M and D as continuous variables. What are the equilibrium concentrations $[M]$ and $[D]$ for $N = 2$ molecules in the cell, assuming these continuous equations and the values above for k_b and k_u ? For $N = 90$ and $N = 10100$ molecules? Numerically solve your differential equation for $N = 2$ and $N = 90$, and verify that your solution settles down to the equilibrium values you found.

We set up differential equations

$$\frac{d[M]}{dt} = 2k_u[D] - 2k_b[M]^2 \quad \frac{d[D]}{dt} = k_b[M]^2 - k_u[D].$$

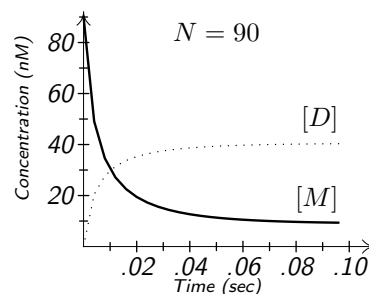
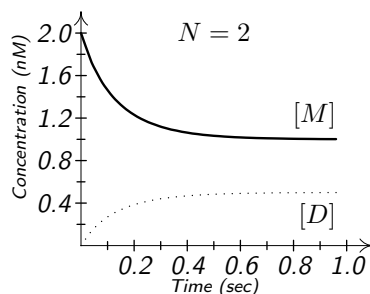
It will be easiest to work with $[M]$, so we use the conservation of number to write $[M] + 2[D] = N$. Substituting out $[D]$ gives

$$\frac{d[M]}{dt} = k_u N - k_u[M] - 2k_b[M]^2.$$

Setting $d[M]/dt = 0$ for equilibrium we solve the quadratic equation, throwing out the (unphysical) negative result,

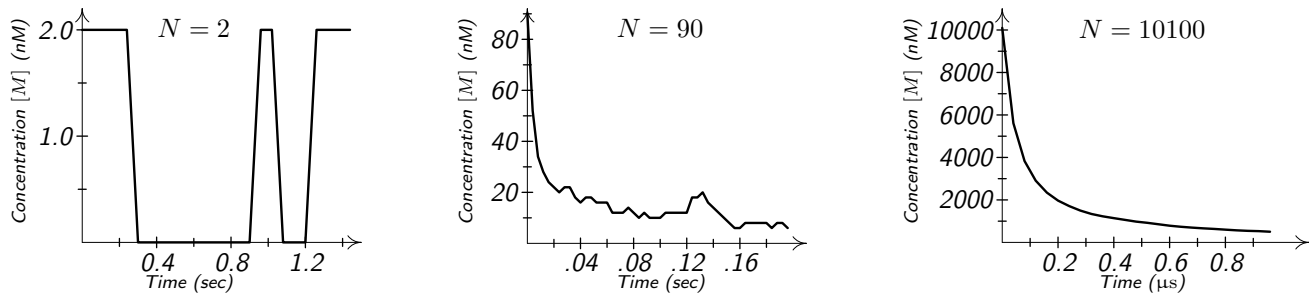
$$[M]_0 = -\frac{k_u}{4k_b} + \sqrt{\frac{k_u}{4k_b} \left(\frac{k_u}{4k_b} + 2N \right)}.$$

For $k_u = 2\text{s}^{-1}$ and $k_b = 1\text{nM}^{-1}\text{s}^{-1}$, we have $k_u/4k_b = \frac{1}{2}\text{nM}$. For $N = 2\text{nM}$ we find $[M]_0 = 1\text{nM}$. For $N = 90\text{nM}$ we get $[M]_0 = 9\text{nM}$. Finally for $N = 10100\text{nM}$ we have $[M]_0 = 100\text{nM}$ and see that indeed 10100 wasn't a typo. Finally, we solve the differential equations numerically (with the initial condition of all monomers) and plot the results below to show agreement.



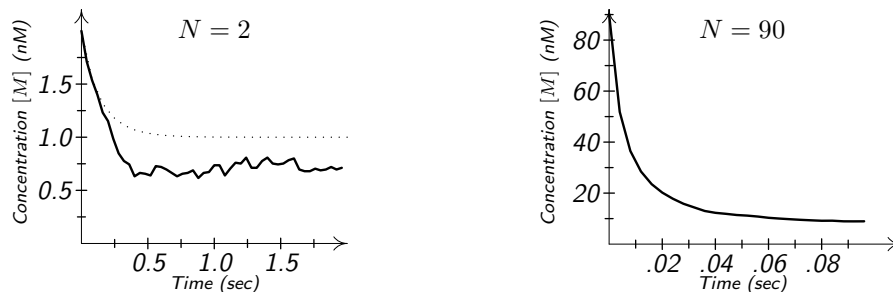
(b) **Stochastic dimerization.** Implement this algorithm for the dimerization reaction of part (a). Simulate for $N = 2$, $N = 90$, and $N = 10100$ and compare a few stochastic realizations with the continuum solution. How large a value of N do you need for the individual reactions to be well described by the continuum equations (say, fluctuations less than $\pm 20\%$ at late times)?

We now try a Monte Carlo simulation. On a side note, the instructions call for a random number distributed with probability $\rho(t) = \Gamma e^{-\Gamma t}$. We can generate this from a uniformly distributed $x \in (0, 1]$ with $t = -\Gamma \log x$. We plot the results below, noting that this method becomes a lot easier than solving the differential equations numerically for very large N . The fluctuations are consistently about ± 5 , but as the numbers get larger this is no longer noticeable. At about $N = 800$, we have late time fluctuations of 20%.



(c) **Average stochastic dimerization.** Find the average of many realizations of your stochastic dimerization in part (b), for $N = 2$ and $N = 90$, and compare with your deterministic solution. How much is the long-term average shifted by the stochastic noise? How large a value of N do you need for the ensemble average of $M(t)$ to be well described by the continuum equations (say, shifted by less than 5% at late times)?

To deal with the large fluctuations from the previous part, we now perform many simulations and average the results. We should expect the fluctuations of the average to decay as $1/\sqrt{T}$ where T is the number of trials. Remarkably this allows us to get reasonable data for $N = 2$. Running enough simulations reveals that the result is consistently shifted from the continuum result by about $-\frac{1}{2}$. This shift drops to 5% when $[M]_0 = 10$, which corresponds to $N = 105$. You should have found pictures similar to the ones below, taken with $T = 250$.



8.12 Entropy Increases! Markov Chains.

(a) Convexity for sums of many terms.

Show by induction that $g(\sum_{\alpha=1}^M \mu_{\alpha} x_{\alpha}) \geq \sum_{\alpha=1}^M \mu_{\alpha} g(x_{\alpha})$, given that $g(x)$ is concave, that $\mu_{\alpha} \geq 0$, $x_{\alpha} \geq 0$ and $\sum_{\alpha=1}^M \mu_{\alpha} = 1$.

The statement is true for $M = 1$. Let's assume it is also true for M . For $M + 1$:

$$\begin{aligned} g\left(\sum_{\alpha=1}^{M+1} \mu_{\alpha} x_{\alpha}\right) &= g\left(\sum_{\alpha=1}^M \mu_{\alpha} x_{\alpha} + \mu_{M+1} x_{M+1}\right) \\ &= g\left((1 - \mu_{M+1}) \sum_{\alpha=1}^M \frac{\mu_{\alpha}}{1 - \mu_{M+1}} x_{\alpha} + \mu_{M+1} x_{M+1}\right) \end{aligned}$$

Clearly, we must have $\sum_{\alpha=1}^M \frac{\mu_{\alpha}}{1 - \mu_{M+1}} = 1$ since $\sum_{\alpha=1}^{M+1} \mu_{\alpha} = 1$. We can now use induction and the definition of concavity:

$$\begin{aligned} g\left(\sum_{\alpha=1}^{M+1} \mu_{\alpha} x_{\alpha}\right) &\geq (1 - \mu_{M+1}) g\left(\sum_{\alpha=1}^M \frac{\mu_{\alpha}}{1 - \mu_{M+1}} x_{\alpha}\right) + \mu_{M+1} g(x_{M+1}) \\ &\geq (1 - \mu_{M+1}) \sum_{\alpha=1}^M \frac{\mu_{\alpha}}{1 - \mu_{M+1}} g(x_{\alpha}) + \mu_{M+1} g(x_{M+1}) \\ &= \sum_{\alpha=1}^{M+1} \mu_{\alpha} g(x_{\alpha}) \end{aligned}$$

Thus, we find that the statement is true for $M + 1$ and thus the statement is true in general by induction.

(b) Show that the free energy decreases for a Markov chain.

Let's define the free energy at step n :

$$F^{(n)} = \sum_{\alpha} \rho_{\alpha}^{(n)} E_{\alpha} + k_B T \rho_{\alpha}^{(n)} \ln(\rho_{\alpha}^{(n)})$$

We have that:

$$\rho_{\alpha}^{(n+1)} = \sum_{\beta} P_{\alpha\beta} \rho_{\beta}^{(n)}$$

Furthermore, let's define the concave function $g(x) = -x \ln(x)$. By using detailed balance and part (a), we have:

$$g\left(\sum_{\beta} P_{\alpha\beta} \rho_{\beta}^{(n)}\right) \geq \sum_{\beta} P_{\alpha\beta} g\left(\frac{\rho_{\alpha}^{(n)}}{\rho_{\beta}^{(n)}} \rho_{\beta}^{(n)}\right)$$

Carefully accounting for the sign in g , we find:

$$F^{(n+1)} \leq \sum_{\alpha\beta} P_{\alpha\beta} \rho_{\beta}^{(n)} \left(E_{\alpha} + k_B T \left(\ln\left(\frac{\rho_{\alpha}^{(n)}}{\rho_{\beta}^{(n)}}\right) + \ln(\rho_{\beta}^{(n)})\right)\right) = F^{(n)}$$

Thus, we find that $F^{(n+1)} \leq F^{(n)}$, which implies that the free energy decreases (and thus the total entropy increases) for a Markov chain.

9.1 Topological Defects in Nematic Liquid Crystals.

Which part of figure 9.23 represents the defect configuration in real space halfway between $S = 1/2$ and $S = -1/2$, corresponding to the intermediate path shown in the middle above?

*A is completely random (temperature beyond that at which defects are meaningful, or else it encloses lots of defects). C is trivial since the order parameter just goes up and down the same side of the hemisphere (doesn't enclose a defect). D is the $S = 1/2$ picture on the left of 9.22, which corresponds to a clockwise traversal about the base of the hemisphere, and E is the $S = -1/2$ counter-clockwise traversal. **B** is the correct answer because it alone has a rotation of 180 degrees about the horizontal axis as we traverse the loop.*

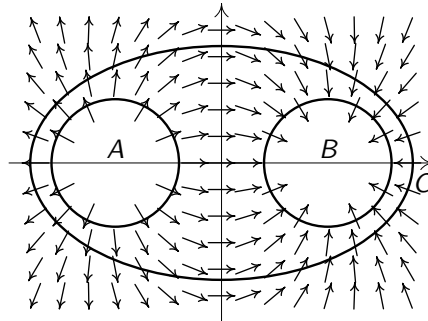
9.2 Topological Defects in the XY Model.

(a) What are the winding numbers of the two defects surrounded by paths A and B in figure 9.20? What should the winding number be around path C, according to the group multiplication law for $\Pi_1(\mathbb{S}^1)$?

A and B both have a winding number of +1. We see that going counter clockwise around the loop results in one full counter clockwise turn in the order parameter. The different defects merely start at different points in this circle. If we rotate every spin 180° , we can continuously transform between them. We can conclude that the winding number about C is $1+1=2$, since C surrounds both A and B defects, and the group "multiplication" in \mathbb{Z}^1 is really addition.

(b) Copy the figure onto a separate sheet of paper, and fill in the region around A and B past C with a smooth, nonsingular, non-vanishing order parameter field of unit vectors.

We smoothly extrapolate the order parameter between A and B past C and find a winding number of +2. The order parameter goes around counterclockwise twice for each counterclockwise loop we make.



(c) Show that the winding number is given by the line integral around the curve D

$$s = \frac{1}{2\pi} \oint \sum_{j=1}^2 (m_1 \partial_j m_2 - m_2 \partial_j m_1) d\ell_j$$

where the two coordinates are x_1 and x_2 , $\partial_j = \frac{\partial}{\partial x_j}$, and ℓ_j is the tangent vector to the contour being integrated around.

We first recognize that $\vec{m} = (\cos \phi, \sin \phi)$ is our unit vector order parameter, so we substitute the components m_1 and m_2 and write everything in terms of ϕ .

$$s = \frac{1}{2\pi} \oint (\cos \phi \partial_j \sin \phi - \sin \phi \partial_j \cos \phi) d\ell_j = \frac{1}{2\pi} \oint (\cos \phi \vec{\nabla} \sin \phi - \sin \phi \vec{\nabla} \cos \phi) \cdot d\vec{\ell}.$$

Now we use the chain rule,

$$\vec{\nabla} f(\phi) = \frac{\partial f}{\partial \phi} \vec{\nabla} \phi.$$

Hence, $\vec{\nabla} \sin \phi = \cos \phi \vec{\nabla} \phi$ and $\vec{\nabla} \cos \phi = -\sin \phi \vec{\nabla} \phi$. Thus,

$$s = \frac{1}{2\pi} \oint (\cos^2 \phi \vec{\nabla} \phi + \sin^2 \phi \vec{\nabla} \phi) \cdot d\vec{\ell}$$

But $\cos^2 \phi + \sin^2 \phi = 1$ so we can now use the Fundamental Theorem of Calculus,

$$s = \frac{1}{2\pi} \oint \vec{\nabla} \phi \cdot d\vec{\ell} = \frac{\Delta \phi}{2\pi}.$$

Around a closed loop, $\Delta \phi$ is a multiple of 2π . This is indeed the definition of the winding number.

9.3 Defect energetics and Total divergence terms.

(a) Show that $\text{curl } \hat{n} = 0$ for the hedgehog. Calculate the free energy of the hedgehog, by integrating the free energy density over the sphere. Compare the free energy to the energy in the same volume with \hat{n} constant.

For the hedgehog, we have $\hat{n} = \hat{r}$. It is easy to see that $\nabla \cdot \hat{r} = \frac{2}{r}$ and $\nabla \times \hat{r} = 0$. Thus, integrating over the volume of the droplet, we have that

$$F_{\text{bulk}} = 8\pi K_{11} R_0.$$

For $\hat{n} = \hat{x}$, the free energy is just 0.

(b) Consider the effects of an additional divergence term. Compare the hedgehog, the anti-hedgehog and the constant vector.

By integrating over the surface, the hedgehog contribution of the divergence term is $4\pi K_0 R_0^2$. Hence, we find that the total free energy in the three cases are given by:

$$\begin{aligned} F_{\text{hedgehog}} &= 8\pi K_{11} R_0 + 4\pi K_0 R_0^2 \\ F_{\text{anti-hedgehog}} &= 8\pi K_{11} R_0 - 4\pi K_0 R_0^2 \\ F_{\hat{x}} &= 0 \end{aligned}$$

For large R_0 , the hedgehog is the ground state for $K_0 < 0$ and the anti-hedgehog is the ground state for $K_0 > 0$.

9.4 Domain Walls in Magnets.

(a) Using these rough estimates find B , minimize the sum, and give a rough value for the energy per unit area of the Bloch wall in terms of K , m_0 , and g .

We consider an Ising magnet in three dimensions, with the boundary condition that $m = \pm m_0$ at $x \rightarrow \pm\infty$. The two terms in the free energy compete, since the first term favors slowly varying $m(x)$ while the second prefers $m(x)$ to be in one well or the other. The barrier $(g/4!)(m^2 - m_0^2)^2$ between the two wells is maximal when $m = 0$, so we write

$$B = \frac{g}{4!}m_0^4.$$

If we assume the magnetization changes linearly from $-m_0$ to $+m_0$ over a width of Δ then the gradient term contributes a free energy density of $\frac{1}{2}K(\nabla m)^2 = 2K(m_0/\Delta)^2$. We have a total free energy per area of the barrier

$$F = \mathcal{F}\Delta = \frac{2K}{\Delta}m_0^2 + \frac{g}{4!}m_0^4\Delta.$$

We minimize this by differentiating $\partial/\partial\Delta$ to find the optimal thickness

$$\Delta = \frac{4}{m_0} \sqrt{\frac{3K}{g}}.$$

We expect this to be the width of the domain wall, up to a multiplicative factor of order one. This gives a rough estimate for the free energy per area

$$F \approx \mathcal{F}\Delta = m_0^3 \sqrt{\frac{gK}{3}}.$$

(b) Find the equation satisfied by that $m(x)$ which minimizes $F = \int \mathcal{F}dx$, given the boundary conditions.

This is a problem in variational calculus, to minimize the functional $F = \int_{-\infty}^{\infty} \mathcal{F}(x)dx$ subject to $m(\pm\infty) = \pm m_0$. We do this just like we would minimize an action by solving the Euler equation. Since we effectively have one dimension, we can change all the $\nabla \rightarrow \frac{\partial}{\partial x}$,

$$\frac{\partial \mathcal{F}}{\partial m} - \frac{d}{dx} \frac{\partial \mathcal{F}}{\partial x'} = 0.$$

This is straightforward to plug in, giving

$$\left(\frac{g}{3!}\right)m(m^2 - m_0^2) - K \frac{\partial^2 m}{\partial x^2} = 0.$$

(c) Show that the solution $m(x)$ has the property that the combination

$$E = (K/2)(\partial m/\partial x)^2 - (g/4!)(m^2 - m_0^2)^2$$

is independent of x .

We differentiate E by x to find

$$\frac{\partial E}{\partial x} = \left(\frac{\partial m}{\partial x} \right) \left[K \frac{\partial^2 m}{\partial x^2} - \left(\frac{g}{3!} \right) m (m^2 - m_0^2) \right] = 0$$

from the result of part (b). Thus the equation for the shape of the domain wall is the same as for of the trajectory of a particle in an inverted double-well potential, with kinetic energy $(K/2)(\partial m/\partial x)^2$ and potential energy $-(g/4!)(m^2 - m_0^2)^2$.

(d) Argue from the boundary conditions that $E = 0$. Using that, find the minimum free energy path $m(x)$ satisfying the boundary conditions $m(\pm\infty) = \pm m_0$. Was your wall thickness estimate of part (a) roughly correct?

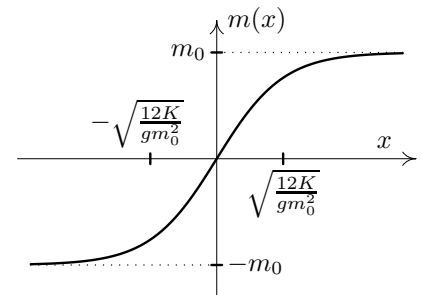
We know m is constant at the boundary, $x \rightarrow \infty$ so $\partial m/\partial x \rightarrow 0$ and $E \rightarrow 0$ at the boundaries. But since E is independent of x , it is identically zero: $E = 0$ for all x . (The particle starts and ends at rest at $x = \pm\infty$.) This gives us a separable differential equation,

$$\frac{dm}{m^2 - m_0^2} = \sqrt{\frac{g}{12K}} dx$$

which we can integrate (recognizing, or else looking in a table to find $d/du(\operatorname{atanh}(u/a)) = a/(a^2 - u^2)$) to get

$$m(x) = m_0 \tanh \left(\sqrt{\frac{g}{12K}} m_0 x \right).$$

Our solution is $m_0 \tanh(x/\Delta)$, with a width $\Delta = \frac{2}{m_0} \sqrt{\frac{3K}{g}}$, a factor of two different from our rough estimate in part (a).



9.5 Landau Theory for the Ising Model.

(a) If there are n spins in the volume ΔV , what value will make $m(\mathbf{x})$ take at temperatures high compared to the interaction J in the Ising model? What values will it take at temperatures very low compared to J ?

At high temperatures $k_B T \gg J$ we expect the spins to be randomly distributed between ± 1 , so that $m(\mathbf{x}) \rightarrow 0$. At low temperatures $k_B T \ll J$ we expect all the spins to align so that the magnetization density is $m(\mathbf{x}) \rightarrow \pm 1$.

(b) What symmetry tells us that the unknown functions A , B , and C do not depend on position \mathbf{x} ? If the magnetization varies on a length scale D , how much smaller would a term involving three derivatives be than the terms B and C that we have kept?

Our free energy density should be translationally invariant, so the unknown functions A , B , and C cannot depend on \mathbf{x} . If the magnetization varies on a (large) length scale D then a term with three derivatives would have one more factor of $1/D$, so that we can ignore it if D is large.

(c) Would the free energy density of equation 9.15 change for a magnet that had a continuous rotational symmetry?

No change. The gradient squared $(\vec{\nabla}m)^2$ and laplacian $(\nabla^2 m)$ are both rotationally invariant already.

(d) Show that the power-law $\beta_{\text{Landau}} = 1/2$ is unchanged in the limit $T \rightarrow T_c$ by keeping additional terms in the Landau expansion ($\frac{h}{6!}m^6$ in equation 9.18). (That is, show that $m(T)/(T - T_c)^\beta$ goes to a constant as $T \rightarrow T_c$.) (Hint: you should get a quadratic equation for m^2 . Keep the root that vanishes at $T = T_c$, and expand in powers of h .) Explore also the alternative, non-generic phase transition where $g \equiv 0$ but $h > 0$: what is β for that transition?

Suppose we keep more terms in the expansion for $A(m^2, T)$. Ignoring all fluctuations, $\vec{\nabla}m \rightarrow 0$ so

$$\mathcal{F}^{\text{ising}} = f_0 + \frac{\mu(T)}{2}m^2 + \frac{g}{4!}m^4 + \frac{h}{6!}m^6 + \dots$$

We differentiate

$$\delta\mathcal{F} = \left(\mu(T)m + \frac{g}{3!}m^3 + \frac{h}{5!}m^5 \right) \delta m = 0.$$

Dividing out the trivial $m = 0$ solution gives

$$\frac{h}{5!}m^4 + \frac{g}{3!}m^2 + \mu(T) = 0$$

which we solve

$$m^2 = -\frac{10g}{h} \pm \sqrt{\left(\frac{10g}{h}\right)^2 - \frac{5!}{h}\mu(T)}.$$

We choose the + solution so that $m \rightarrow 0$ as $T \rightarrow T_c$ and expand to lowest order in $\mu(T)$,

$$m^2 = \frac{10g}{h} \left(-1 + 1 - \left(\frac{h}{10g}\right)^2 \frac{1}{2} \frac{5!}{h} \mu(T) \right) = -\frac{6a(T - T_c)}{g}.$$

This is the same result as above, $m \sim \pm\sqrt{T - T_c}$, $\beta = \frac{1}{2}$, to leading order in $T - T_c$. We could expand to higher orders in h and μ which may give different behavior for a time, but ultimately as $T \rightarrow T_c$ only this term remains.

Now we consider the qualitatively different transition when $g \equiv 0$. Backtracking we have

$$\mu(T) + \frac{h}{5!}m^4 = 0$$

so that $m \sim \pm(T - T_c)^{1/4}$, or $\beta = 1/4$.

The reason these exponents are wrong is the fact that we have ignored the thermal fluctuations on length scales larger than the coarse-graining length we used to define our free energy density \mathcal{F} . (Minimizing the free energy, rather than sampling it with a thermal distribution, is ignoring the thermal fluctuations.) We shall see in Chapter 12 that near continuous phase transitions there are fluctuations on all length scales, from whose effects the true critical exponent β emerges.

9.6 Symmetries and Wave Equations .

(a) What common equation of motion results from the choice $G(a_1, a_2, \dots) = a_3 - Da_4$?

The diffusion equation!

(b) Give the missing terms, multiplying functions j, k, \dots, m .

The terms are:

$$\left(\frac{\partial y}{\partial x}\right)^2, \frac{\partial y}{\partial t} \frac{\partial y}{\partial x}, \frac{\partial^2 y}{\partial t^2}, \frac{\partial^2 y}{\partial x^2}$$

(c) Which three terms are left after imposing the symmetries?

The terms are:

$$\frac{\partial y}{\partial t}, \frac{\partial^2 y}{\partial t^2}, \frac{\partial^2 y}{\partial x^2}$$

with $\frac{\partial y}{\partial t}$ not part of the wave equation.

(d) Show that the third term is not invariant under boosts. Show that Kelvin damping behaves like the other terms in the wave equation.

applying differentiation trivially yields the result.

9.7 Superfluid Order and Vortices.

(a) If N non-interacting bosons reside in a state $\chi(\vec{r})$, write an expression for the net current density $J(\vec{r})$. Write the complex field $\chi(\vec{r})$ in terms of the amplitude and a phase, $\chi(\vec{r}) = |\chi(\vec{r})| \exp(i\phi(\vec{r}))$. We write the superfluid density as $n_s = N|\chi|^2$. Give the current J in terms of ϕ and n_s . What is the resulting superfluid velocity $v = J/n_s$?

We can use the standard single-particle expression for current density, multiplied by N (if you don't believe me, you can work it out for the N -particle state $\Psi = \prod_n \chi(\vec{r}_n)$ and $\vec{J}(\vec{r}) = \frac{i\hbar}{2m} \int d^3\vec{r}_1 \dots d^3\vec{r}_n \sum_\ell \delta(\vec{r} - \vec{r}_\ell) (\Psi \vec{\nabla}_\ell \Psi^ - \text{c.c.})$ and get the same factor of N),*

$$\vec{J}(\vec{r}) = \frac{i\hbar}{2m} N \left(\chi \vec{\nabla} \chi^* - \chi^* \vec{\nabla} \chi \right).$$

Plugging in $\chi(\vec{r}) = |\chi(\vec{r})| \exp(i\phi(\vec{r}))$, we find $\vec{J} = (N\hbar/m)|\chi|^2 \vec{\nabla} \phi$. With $n_s = N|\chi|^2$, this is $\vec{J} = (\hbar n_s/m) \vec{\nabla} \phi$, so the superfluid velocity is

$$\vec{v} = \frac{\vec{J}}{n_s} = \frac{\hbar}{m} \vec{\nabla} \phi.$$

(b) Consider a torus filled with an ideal Bose gas at $T = 0$ with the hole along the vertical axis: the superfluid is condensed into a state which is rotating around the hole. Using your formula from part (a) and the fact that $\phi + 2n\pi$ is indistinguishable from ϕ for any integer n , show that the circulation $\oint \vec{v} \cdot d\vec{r}$ around the hole is quantized. What is the quantum of circulation.

We integrate around a closed loop, say from $\varphi = 0$ to $\varphi = 2\pi$,

$$\oint \vec{v} \cdot d\vec{\ell} = \frac{\hbar}{m} \int_{\varphi=0}^{\varphi=2\pi} \vec{\nabla} \phi \cdot d\vec{\ell} = \frac{\hbar}{m} (\phi(2\pi) - \phi(0)),$$

where we use the Fundamental Theorem of Calculus again. Since $\phi(2\pi)$ and $\phi(0)$ are actually evaluated at the same point, we conclude that $\phi(2\pi) = \phi(0) + 2n\pi$. Therefore the circulation is

$$\oint \vec{v} \cdot d\vec{\ell} = \frac{n\hbar}{m}$$

for $n \in \mathbb{Z}$ and the quantum of circulation is \hbar/m .

(c) Treat $\phi(\vec{r})$, the phase of the superconducting wave function, as the topological order parameter of the superfluid. Is the order parameter a closed loop? Classify the type of vortex lines in a superfluid.

Since changing the phase ϕ by 2π leaves the wave function unchanged, we can consider the space of possible ϕ values to be the line segment $[0, 2\pi)$ with periodic boundary conditions. This is topologically equivalent to a circle, so the order parameter space is a circle. The ‘defects’ in a superfluid are like little vortices, which can be characterized by the value of the circulation on a path encircling them. The possible values of the circulation are specified by an integer n , so the first Homotopy group of the order parameter space is the integers, \mathbb{Z} .

9.8 Superfluids: Density Matrices and ODLRO.

(a) What does the reduced...

At a sufficiently low temperature so that all the bosons have condensed into a normalized state $\chi(\mathbf{r})$, the N -particle wave function is just $\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = \prod_i \chi(\mathbf{r}_i)$. The reduced density matrix for this state is:

$$\hat{\rho}_2(\mathbf{r}', \mathbf{r}) = N \int d^3r_2 \dots d^3r_N \chi^*(\mathbf{r}') \chi(\mathbf{r}) \prod_{i=2}^N \chi^*(\mathbf{r}_i) \chi(\mathbf{r}_i).$$

Since $\chi(\mathbf{r}_i)$ is normalized, this is:

$$\hat{\rho}_2(\mathbf{r}', \mathbf{r}) = N \chi^*(\mathbf{r}') \chi(\mathbf{r}).$$

(b) Show that the ket...

In the expression for the ket $|\Psi\rangle$, the function $\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N)$ is precisely the wave function in the coordinate representation: $\langle \mathbf{r}_1 \dots \mathbf{r}_N | \Psi \rangle$. This is important to note because a many-particle wave function of identical bosons must be totally symmetric under interchange of any two particles. To show that $|\Psi\rangle$ is normalized, proceed as follows:

$$\langle \Psi | \Psi \rangle = \frac{1}{N!} \int d^3 r'_1 \dots d^3 r'_N d^3 r_1 \dots d^3 r_N \Psi^*(\mathbf{r}'_i) \Psi(\mathbf{r}_i) \langle 0 | a(\mathbf{r}'_N) \dots a(\mathbf{r}'_1) a^\dagger(\mathbf{r}_1) \dots a^\dagger(\mathbf{r}_N) | 0 \rangle.$$

Now focus on the a 's and a^\dagger 's:

$$a(\mathbf{r}'_1) a^\dagger(\mathbf{r}_1) = \delta(\mathbf{r}_1 - \mathbf{r}'_1) + a^\dagger(\mathbf{r}_1) a(\mathbf{r}'_1).$$

We commute $a(\mathbf{r}'_1)$ through to the right, each time getting a delta function $\delta(\mathbf{r}'_1 - \mathbf{r}_i)$ times a term with \mathbf{r}_i and \mathbf{r}'_1 absent, until $a(\mathbf{r}'_1)$ is all the way to the right and annihilates on $|0\rangle$. This gives us a sum of N terms. Next we repeat the procedure with $a(\mathbf{r}'_2)$, so each of the N terms expands into $N - 1$, giving a total of $N(N - 1)$ terms each containing two delta functions: $\delta(\mathbf{r}'_1 - \mathbf{r}_i) \delta(\mathbf{r}'_2 - \mathbf{r}_j)$. Continuing until there are no a 's left to commute through, we end up with $N!$ terms each containing N delta functions:

$$\delta(\mathbf{r}'_1 - \mathbf{r}_{j_1}) \delta(\mathbf{r}'_2 - \mathbf{r}_{j_2}) \dots \delta(\mathbf{r}'_N - \mathbf{r}_{j_N}),$$

where $\{j_1, \dots, j_N\}$ is a permutation of $\{1, \dots, N\}$. In fact, the $N!$ terms correspond exactly to the $N!$ permutations of $\{1, \dots, N\}$. At any rate, under the integral we next rename the variables in each term so the delta's look the same, and then we find that because of the symmetry of Ψ each term is in fact identical. The delta's kill N of the integrals, leaving:

$$\langle \Psi | \Psi \rangle = \frac{1}{N!} \sum_{\text{permutations}} \int d^3 r_1 \dots d^3 r_N \Psi^* \Psi = \frac{1}{N!} \sum (1) = 1,$$

where we have used the fact that $\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N)$ is normalized.

Showing that $\langle \Psi | a^\dagger(\mathbf{r}') a(\mathbf{r}) | \Psi \rangle$ is the reduced density matrix for the pure state $|\Psi\rangle$ involves a similar procedure. First write it out explicitly:

$$\langle \Psi | a^\dagger(\mathbf{r}') a(\mathbf{r}) | \Psi \rangle = \frac{1}{N!} \int d^3 r'_1 \dots d^3 r'_N d^3 r_1 \dots d^3 r_N \Psi^*(\{\mathbf{r}'_i\}) \Psi(\{\mathbf{r}_i\})^*$$

$$\langle 0 | a(\mathbf{r}'_N) \dots a(\mathbf{r}'_1) a^\dagger(\mathbf{r}') a(\mathbf{r}) a^\dagger(\mathbf{r}_1) \dots a^\dagger(\mathbf{r}_N) | 0 \rangle.$$

Now commute $a(\mathbf{r})$ through to the right to get N terms each containing $\delta(\mathbf{r} - \mathbf{r}_i)$. In each of these, commute $a^\dagger(\mathbf{r}')$ through to the left to get N more terms containing $\delta(\mathbf{r}' - \mathbf{r}'_j)$. This gives us a total of N^2 terms so far, each containing two delta functions. We get $(N - 1)!$ more terms from permuting the remaining a 's through, giving us a grand total of $N^2(N - 1)!$ terms, each with $N + 1$ delta functions. We now use again the fact that $\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N)$ is symmetric under interchange of any two particles to rename variables in each term until we have, after integrating away $N - 1$ of the delta functions:

$$\begin{aligned} \langle \Psi | a^\dagger(\mathbf{r}') a(\mathbf{r}) | \Psi \rangle &= \frac{1}{N!} N^2 (N - 1)! \int d^3 r'_1 d^3 r_1 \dots d^3 r_N \delta(\mathbf{r} - \mathbf{r}_1) \delta(\mathbf{r}' - \mathbf{r}'_1) \Psi^*(\mathbf{r}'_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \\ &= N \int d^3 r_2 \dots d^3 r_N \Psi^*(\mathbf{r}', \mathbf{r}_2, \dots, \mathbf{r}_N) \Psi(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N) = \hat{\rho}_2(\mathbf{r}', \mathbf{r}). \quad \text{QED.} \end{aligned}$$

So the reduced density matrix is equal to the inner product or overlap of $a(\mathbf{r})|\Psi\rangle$ with $a(\mathbf{r}')|\Psi\rangle$. $|\Psi\rangle$ is an N -particle state, so these are $(N - 1)$ -particle states. If removing a particle at \mathbf{r} affected the wave function only in a local neighborhood, we would expect this overlap to be zero for \mathbf{r} far from \mathbf{r}' . When it's non-zero even as $|\mathbf{r} - \mathbf{r}'| \rightarrow \infty$, as in part (c) below, it means that removing a particle at a particular point significantly affects the wave function everywhere: the particles in a

superfluid exist in non-localized states—this is the long range order.

(c) What is ...

The hint in the more recent versions of the text was quite helpful with this part. Assume we know the value of $\hat{\rho}_2(\mathbf{r}', \mathbf{r})$ everywhere. Let the points \mathbf{a} , \mathbf{b} , and \mathbf{c} be far from each other, so that $\hat{\rho}_2(\mathbf{r}', \mathbf{r}) \rightarrow \Psi^*(\mathbf{r}')\Psi(\mathbf{r})$ applies. Then:

$$|\Psi^*(\mathbf{a})| = \left(\frac{\hat{\rho}_2(\mathbf{a}, \mathbf{c})\hat{\rho}_2(\mathbf{b}, \mathbf{a})}{\hat{\rho}_2(\mathbf{b}, \mathbf{c})} \right)^{1/2}.$$

Thus for any position \mathbf{r} far from \mathbf{a} , $\hat{\rho}_2(\mathbf{a}, \mathbf{r}) = \Psi^*(\mathbf{a})\Psi(\mathbf{r})$ determines $\Psi(\mathbf{r})$ up to a phase. For points \mathbf{r} near \mathbf{a} , we pick a point \mathbf{e} , say, that is far from \mathbf{r} and \mathbf{a} , and then:

$$\Psi(\mathbf{r}) = \Psi(\mathbf{a}) \frac{\hat{\rho}_2(\mathbf{e}, \mathbf{r})}{\hat{\rho}_2(\mathbf{e}, \mathbf{a})},$$

determining $\Psi(\mathbf{r})$ up to the same phase as before. Thus, knowing the reduced density matrix everywhere allows us to determine $\Psi(\mathbf{r})$ anywhere up to an overall constant phase.

For the non-interacting Bose condensate of part (a), we saw that $\hat{\rho}_2(\mathbf{r}', \mathbf{r}) = N\chi^*(\mathbf{r}')\chi(\mathbf{r})$. Hence, $\Psi(\mathbf{r}) = \sqrt{N}e^{i\phi_0}\chi(\mathbf{r})$, where ϕ_0 is an arbitrary phase.

(d) Consider a zero-...

The wave function of the N -particle condensate is $\Psi = \prod_n \chi(\mathbf{r}_n) = \prod_n (|\chi(\mathbf{r}_n)|e^{i\phi}) = |\chi|^N e^{iN\phi}$. The phase of the N -particle state is $N\phi$.

(e) Consider the Hamiltonian...

If \mathcal{H} conserves total particle number, then each term in \mathcal{H} must contain an equal number of a 's and a^\dagger 's. Multiplying the creation operators by $e^{i\zeta}$ and the annihilation operators by $e^{-i\zeta}$ therefore leaves \mathcal{H} unchanged, as the exponents all cancel out.

(f) Show that this...

What does it mean for the Hamiltonian to be invariant? Certainly not that $\mathcal{H}(p - \hbar\nabla\zeta) = \mathcal{H}(p)$. The Hamiltonian is the operator for the energy, and the expectation value of the energy of a state $|\psi\rangle$ is $\langle\psi|\mathcal{H}|\psi\rangle$. For \mathcal{H} to be invariant means that this expectation value is unchanged by the transformation:

$$\langle\tilde{\psi}|\tilde{\mathcal{H}}|\tilde{\psi}\rangle = \langle\psi|e^{-i\zeta}\mathcal{H}(p - \hbar\nabla\zeta)e^{i\zeta}|\psi\rangle = \langle\psi|\mathcal{H}(p)|\psi\rangle.$$

This will hold for any state $|\psi\rangle$ if and only if:

$$e^{-i\zeta} \mathcal{H}(p - \hbar \nabla \zeta) e^{i\zeta} = \mathcal{H}(p).$$

It is straightforward to show that this is in fact the case for the single-particle Hamiltonian $\mathcal{H}(p) = p^2/2m = (\frac{\hbar}{i}\nabla)^2/2m$.

(g) Consider the Hamiltonian...

We want to show that $e^{-i(q/e)\zeta} \mathcal{H}(\mathbf{A}') e^{i(q/e)\zeta} = \mathcal{H}(\mathbf{A})$ for a suitable gauge transformation $\mathbf{A}' \rightarrow \mathbf{A} + \mathbf{W}$, with \mathbf{W} to be determined:

$$\begin{aligned} e^{-i(q/e)\zeta} \mathcal{H}(\mathbf{A}') e^{i(q/e)\zeta} &= e^{-i(q/e)\zeta} \frac{1}{2m} \left(\frac{\hbar}{i} \nabla - \frac{q}{c} \mathbf{A}' \right)^2 e^{i(q/e)\zeta} \\ &= \frac{1}{2m} \left[\frac{\hbar^2 q^2}{e^2} (\nabla \zeta)^2 - \frac{iq\hbar^2}{e} \nabla^2 \zeta - \frac{2iq\hbar^2}{e} \nabla \zeta \cdot \nabla \right. \\ &\quad \left. - \hbar^2 \nabla^2 + \frac{i\hbar q}{c} \nabla \cdot \mathbf{A}' - \frac{2\hbar q^2}{ce} \mathbf{A}' \cdot \nabla \zeta + \frac{2iq\hbar}{c} \mathbf{A}' \cdot \nabla + \frac{q^2}{c^2} \mathbf{A}'^2 \right]. \end{aligned}$$

Putting in $\mathbf{A}' = \mathbf{A} + \mathbf{W}$ and setting the above expression equal to $\mathcal{H}(\mathbf{A})$ yields:

$$\begin{aligned} \frac{\hbar^2 q^2}{e^2} (\nabla \zeta)^2 - \frac{iq\hbar^2}{e} \nabla^2 \zeta - \frac{2iq\hbar^2}{e} \nabla \zeta \cdot \nabla + \frac{i\hbar q}{c} \nabla \cdot \mathbf{W} - \frac{2\hbar q^2}{ce} \mathbf{A} \cdot \nabla \zeta \\ - \frac{2\hbar q^2}{ce} \mathbf{W} \cdot \nabla \zeta + \frac{2iq\hbar}{c} \mathbf{W} \cdot \nabla + \frac{q^2}{c^2} \mathbf{W}^2 + \frac{2q^2}{c^2} \mathbf{A} \cdot \mathbf{W} = 0. \end{aligned} \quad (*)$$

This must hold for any \mathbf{A} , so we set the terms with \mathbf{A} separately to zero:

$$\mathbf{A} \cdot \left(\frac{2q^2}{c^2} \mathbf{W} - \frac{2\hbar q^2}{ce} \nabla \zeta \right) = 0$$

$$\iff \mathbf{W} = \frac{\hbar c}{e} \nabla \zeta.$$

And we can check that this choice of \mathbf{W} does indeed satisfy (*), so $\mathbf{A}' = \mathbf{A} + \frac{\hbar c}{e} \nabla \zeta$ is the required gauge transformation.

10.1 Microwave Background Radiation.

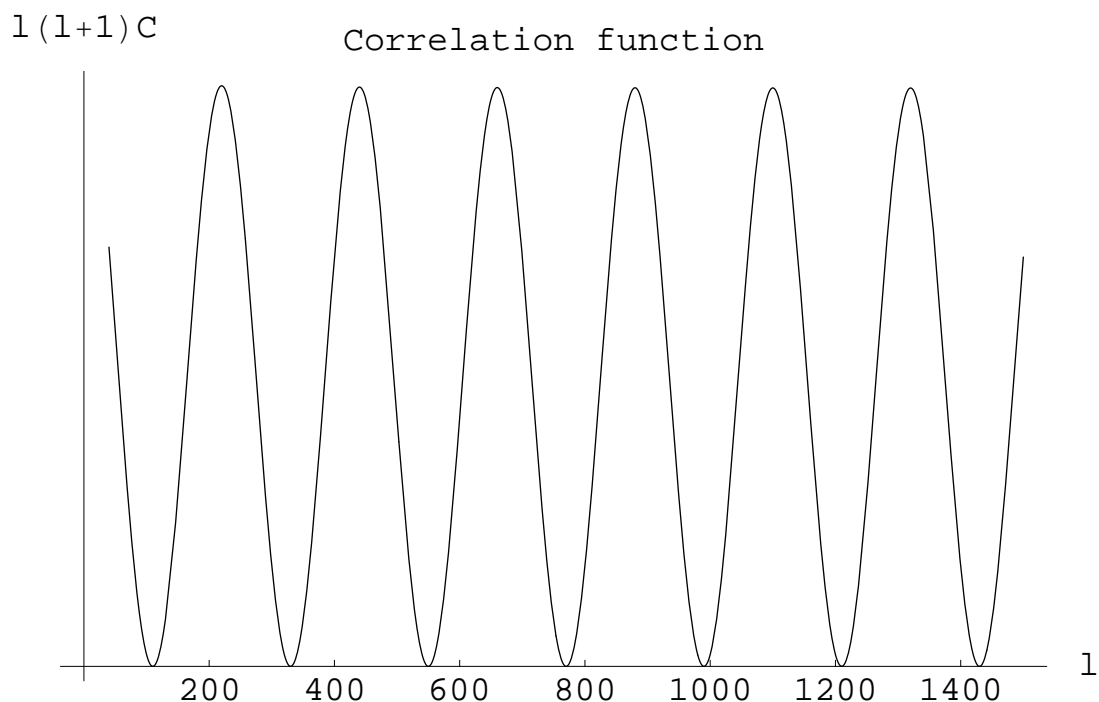
(a) Given an initial $\Theta(k, t=0)$ and $\Theta_t(k, t=0) = 0$, calculate $\Theta(k, t)$, \hat{C} and the value of L that insures the first peak is at $l = 220$.

Using the wave equation for one fourier mode, we obtain:

$$(1 + R)\Theta_{tt} = -\frac{1}{3}c^2 k^2 \Theta$$

This can be solved to yield:

$$\Theta(k, t) = \Theta_0 \cos \left(\sqrt{\frac{c^2 k^2}{3(1+R)}} t \right)$$



and thus:

$$C(\hat{k}, t) = \frac{A}{k^2} \cos^2 \left(\sqrt{\frac{c^2 k^2}{3(1+R)}} t \right)$$

demanding that this peaks when $l_0 = 220$ at t_0 (the decoupling time) implies:

$$\sqrt{\frac{c^2 k_0^2}{3(1+R)}} t_0 = \pi$$

$$L = \frac{2ct_0 l_0}{\sqrt{3(1+R)}} = 6.98 \times 10^{25} \text{ cm} = 7.4 \times 10^7 \text{ LYrs}$$

(b) Plot the correlation function

As can be seen below, the calculated function barely looks like plot 10.14. It does oscillate at a similar frequency...

10.2 Pair distributions and molecular dynamics.

(a) Find analytic expressions for the density, the radial distribution function and the potential energy.

We have:

$$\rho(x) = \left\langle \sum_{i=1}^N \delta(x - x_i) \right\rangle = N \langle \delta(x - x_1) \rangle$$

since all the particles are equivalent. Since the system is homogeneous, each particle has the same probability to be anywhere in the sample. This means that the probability distribution must be of the form $P(x_1) = \frac{1}{V}$. Thus:

$$\rho(x) = N \int_V \frac{1}{V} \delta(x - x_1) d^3x_1 = \frac{N}{V}$$

Next, the correlation function is given by $C(x, x') = \langle \sum_{i \neq j} \delta(x - x_i) \delta(x' - x_j) \rangle$. Since the system is homogeneous, the center of mass of the particles x_i and x_j can be anywhere with equal probability. Thus, $\hat{P}(x_i, x_j) = P_{cm}(\frac{x_i + x_j}{2}) P(x_i - x_j) = \frac{1}{V} P(x_i - x_j)$. Define $r = x - x'$. Then:

$$C(x, x') = C(r) = \sum_{i \neq j} \frac{1}{V} \int_V \int_V \delta(r + x' - x_i) \delta(x' - x_j) P(x_i - x_j) = \sum_{i \neq j} \frac{1}{V} \int_V P(r_{ij}) \delta(r - r_{ij}) = \frac{1}{V} \langle \sum_{i \neq j} \delta(r - r_{ij}) \rangle$$

This implies that:

$$g(r) = \frac{C(r)}{\rho^2} = \langle \frac{V}{N^2} \sum_{i \neq j} \delta(r - r_{ij}) \rangle$$

If the system is uncorrelated, then $P(r_{ij}) = \frac{1}{V}$ and $g(r) \equiv 1$. The average energy is given by:

$$\langle E \rangle = \frac{1}{2} \langle \sum_{i \neq j} E(r_{ij}) \rangle = \frac{N}{2} \frac{N}{V} \int_V d^3r g(r) E(r)$$

.

(b) Find a relation between $g(r)$ and $h(r)$.

The average number of particles separated by a distance $r_n < r < r_n + \Delta$ is given by:

$$h(r_n) = \frac{1}{2} \int_V d^3r' \int_{r_n}^{r_n + \Delta} 4\pi r'^2 c(r) \simeq \frac{N^2}{2V} 4\pi r_n^2 g(r_n) \Delta$$

. Thus:

$$g(r) = \frac{2V}{N^2} \frac{h(r)}{4\pi r^2 \Delta r}$$

.

In two dimensions, $V \rightarrow A$ and $4\pi r^2 \rightarrow 2\pi r$, hence:

$$g_{2D}(r) = \frac{2A}{N^2} \frac{h(r)}{2\pi r \Delta r}$$

.

(c) What is the minimum of the Lennard-Jones potential, and where does it occur?

The pair potential is given by:

$$E(r) = 4 \left(\frac{1}{r^{12}} - \frac{1}{r^6} \right).$$

Straightforward calculus yields:

$$E_{min} = -1; r_{min} = 2^{\frac{1}{6}}.$$

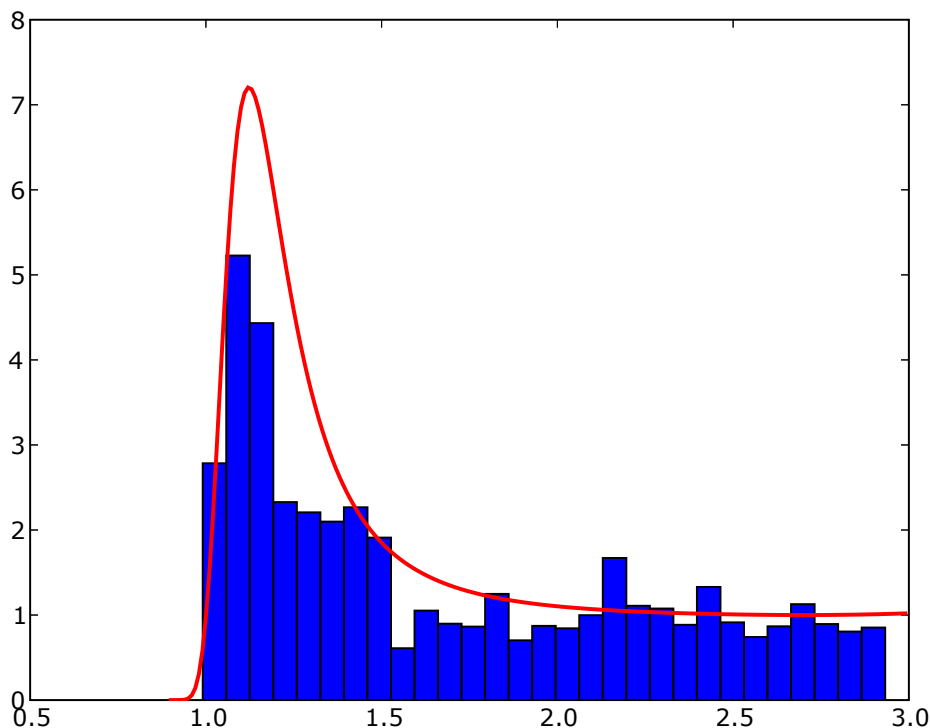


FIG. 32: Gas radial distribution function along with theoretical expression

(d) (e) Gas radial distribution function.

See figure 32. The theoretical expression represents the data well. We expect even better agreement at higher temperature or lower densities. We see that the maximum of the radial distribution function is at the position of the minimum of the pair potential, as expected.

(f) (g) Liquid radial distribution function.

See figure 33. The radial distribution function peaks at the distance of the nearest and second nearest neighbor. The average number of neighbors

(h) Solid radial distribution function.

See figure 34. The fluctuations are much smaller than the lattice constant.

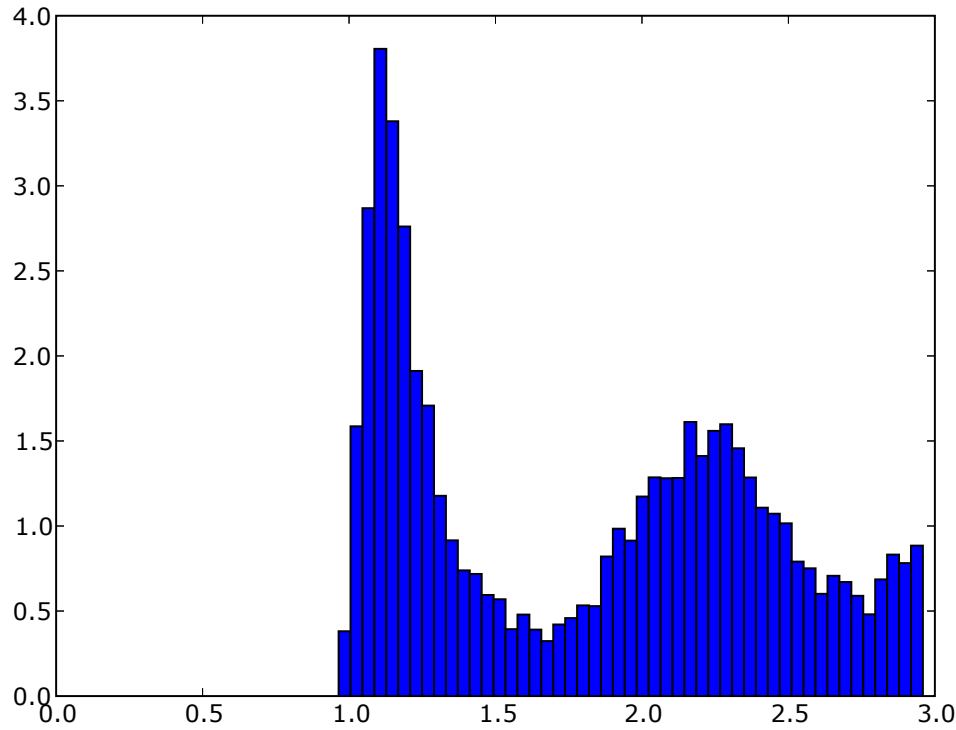


FIG. 33: Liquid radial distribution function

10.3 Damped Oscillators.

(a) **Susceptibility.** Find the AC susceptibility $\tilde{\chi}(\omega)$ for the oscillator. Plot χ' and χ'' for $\omega = m = 1$ and $\gamma = 0.2, 2$, and 5.

We start by making clear our convention for Fourier transforms. We define

$$\tilde{f}(\omega) = \int_{-\infty}^{\infty} dt f(t) e^{i\omega t} \quad f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \tilde{f}(\omega) e^{-i\omega t}.$$

We see from this definition that $d/dt \rightarrow -i\omega$. The susceptibility is gotten by taking $f(t) \rightarrow \delta(t)$ and $\theta(t) \rightarrow \chi(t)$. We Fourier transform the equation of motion,

$$-\omega^2 \tilde{\chi}(\omega) = -\omega_0^2 \tilde{\chi}(\omega) + i\gamma \tilde{\chi}(\omega) + \frac{1}{m} \tilde{f}(\omega).$$

We can now solve algebraically to find $\tilde{\chi}(\omega) = \frac{1}{m} \frac{1}{\omega_0^2 - \omega^2 - i\gamma\omega} = \chi' + i\chi''$. Completing the square we have

$$\chi'(\omega) = \frac{1}{m} \frac{\omega_0^2 - \omega^2}{(\omega_0^2 - \omega^2)^2 + \gamma^2 \omega^2} \quad \chi''(\omega) = \frac{1}{m} \frac{\gamma\omega}{(\omega_0^2 - \omega^2)^2 + \gamma^2 \omega^2}.$$

We plot this below

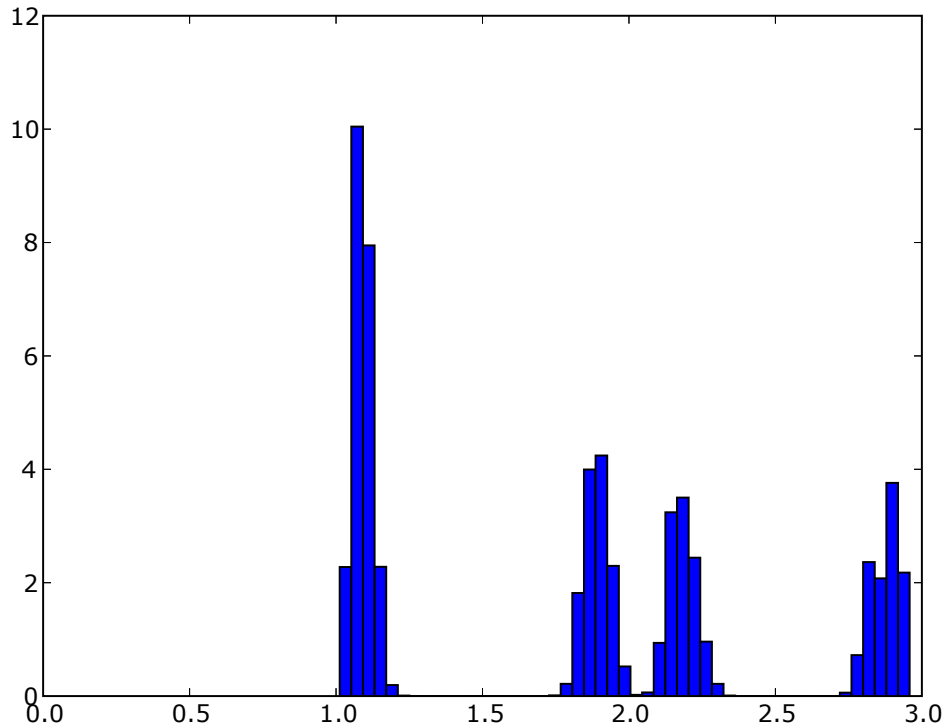
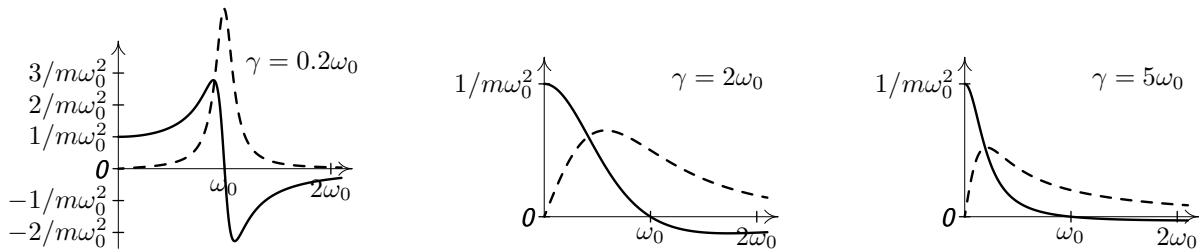


FIG. 34: Solid radial distribution function



The solid line is χ' and the dashed line is χ'' . We observe that these correspond to underdamping, critical damping, and overdamping respectively.

(b) **Causality and Critical Damping.** Check, for positive damping γ , that your $\chi(\omega)$ is causal by examining the singularities in the complex ω plane. At what value of γ do the poles begin to sit on the imaginary axis?

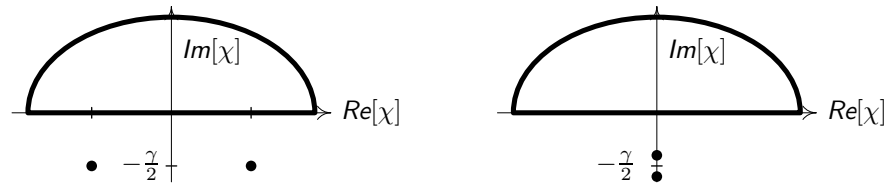
We take $\gamma > 0$. Now we compute $\chi(t)$ for $t < 0$ to test causality. We write this as an inverse Fourier transform,

$$\chi(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \tilde{\chi}(\omega) e^{-i\omega t}.$$

This is analytic, so we look at the poles. $\tilde{\chi}$ blows up when its denominator $\omega_0^2 - \omega^2 - i\gamma\omega = 0$. We solve this quadratic to find

$$\omega = -\frac{i\gamma}{2} \pm \sqrt{\omega_0^2 - \left(\frac{\gamma}{2}\right)^2}.$$

Fortunately, since $t < 0$ we are integrating around the upper half-plane, so that $e^{-i\omega t} \rightarrow 0$. Thus, when $\gamma < 2\omega_0$ we are underdamped and the poles are in the third and fourth quadrants. When $\gamma > 2\omega_0$ we are overdamped and the poles are on the negative imaginary axis. The poles are never in the upper half-plane, so our integral is always zero and χ is causal.



(c) **Dissipation and Susceptibility.** Given a forcing $f(t) = A \cos(\omega t)$, solve the equation and calculate $\theta(t)$. Calculate the average power dissipated by integrating your resulting formula for $f \, d\theta/dt$. Do your answers for the power and χ'' agree with the general formula for power dissipation, equation 10.35?

We now find $\theta(t)$ for $f(t) = A \cos(\omega t)$. We can use our susceptibility to do this, since $\tilde{\theta}(\omega') = \tilde{\chi}(\omega') \tilde{f}(\omega')$. Writing $\cos(\omega t)$ as $\frac{1}{2}(e^{i\omega t} + e^{-i\omega t})$ we find

$$\tilde{f}(\omega') = \frac{A}{2} \int_{-\infty}^{\infty} (e^{i(\omega' + \omega)t} + e^{i(\omega' - \omega)t}) dt = A\pi\delta(\omega' + \omega) + A\pi\delta(\omega' - \omega).$$

Therefore

$$\theta(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega' e^{-i\omega' t} \tilde{\chi}(\omega') \tilde{f}(\omega') = A\chi'(\omega) \cos \omega t + A\chi''(\omega) \sin \omega t = \text{Re} [A\tilde{\chi}(\omega) e^{-i\omega t}].$$

Now we differentiate $\theta(t)$ to find the power $f\dot{\theta} = (A \cos \omega t) A\omega(-\chi'(\omega) \sin \omega t + \chi''(\omega) \cos \omega t)$. We average over a cycle

$$p(\omega) = \langle \dot{\theta} f \rangle = -\langle A^2 \omega \chi'(\omega) \sin \omega t \cos \omega t \rangle + \langle A^2 \omega \chi''(\omega) \cos^2 \omega t \rangle = \frac{1}{2} A^2 \omega \chi''(\omega).$$

This agrees with 10.65 in the text.

(d) **Correlations and Thermal Equilibrium.** Use the fluctuation-dissipation theorem to calculate the correlation function $\tilde{C}(\omega)$ from $\chi''(\omega)$, where

$$C(t - t') = \langle \theta(t) \theta(t') \rangle.$$

Find the equal-time correlation function $C(0) = \langle \theta^2 \rangle$, and show that it satisfies the equipartition theorem.

The fluctuation-dissipation theorem gives $\tilde{C}(\omega) = \frac{2}{\beta\omega} \chi''(\omega)$. We therefore have

$$C(t) = \frac{1}{2\pi} \frac{2\gamma}{m\beta} \int_{-\infty}^{\infty} \frac{e^{-i\omega t} d\omega}{(\omega_0^2 - \omega^2)^2 + \gamma^2 \omega^2}.$$

This integral is not nearly as straightforward as the text suggests and the integral in the hint is useless (unless I missed something). We must therefore do a contour integral. If we take $t = 0$, this is not feasible, but we will instead take $t \rightarrow 0^-$ so that we can integrate over the upper half-plane (the opposite would work just as well). The denominator is a fourth order polynomial in ω so we will have four poles which we can find by factoring. First we factor a complex difference of squares

$$(\omega_0^2 - \omega^2)^2 + \gamma^2 \omega^2 = (\omega_0^2 - \omega^2 + i\gamma\omega)(\omega_0^2 - \omega^2 - i\gamma\omega) = (\omega^2 - i\gamma\omega - \omega_0^2)(\omega^2 + i\gamma\omega - \omega_0^2).$$

Now we see that *only the first term vanishes in the upper half-plane giving two poles. We use the residues from the poles at $\omega = i\gamma/2 \pm \sqrt{\omega_0^2 - (\gamma/2)^2}$ and multiply by $2i\pi$ to get (after some work)*

$$\int_{-\infty}^{\infty} \frac{e^{-i\omega t} d\omega}{(\omega_0^2 - \omega)^2 + \gamma^2 \omega^2} = \frac{\pi}{\gamma \omega_0^2} \quad \text{as } t \rightarrow 0.$$

We therefore have $C(0) = \langle \theta^2 \rangle = k_B T / m \omega_0^2$. Recalling that $V(\theta) = \frac{1}{2} m \omega_0^2 \theta^2$ we see that the equipartition theorem holds:

$$\langle V(\theta) \rangle = \frac{1}{2} m \omega_0^2 \langle \theta^2 \rangle = \frac{1}{2} k_B T.$$

10.4 Spin.

[This is the answer to an older version of the exercise. The exercise may have been improved since this was written.] There was some confusion about units and factors of $g\mu_b$ in this problem, so first let me digress with a somewhat long, rambling discussion of these things...

Notice that the usual interaction energy for a magnetic dipole is $E = -\vec{\mu} \cdot \vec{H}$, so the spin we have in this problem is like an electron, having a magnetic moment anti-parallel to its spin and consequently an energy $E = +g\mu_b \vec{s} \cdot \vec{H}$ (assuming $g > 0$). The governing dynamical equation given in the problem is:

$$\frac{d\langle S(t) \rangle}{dt} = \Gamma \langle S(t) \rangle + \Gamma \chi_0 H(t).$$

As defined by this equation, χ_0 is a static ‘spin susceptibility’ $\chi_0 \equiv \frac{\partial \langle S \rangle}{\partial H}$. There are two other susceptibilities that are relevant. The ‘magnetic susceptibility’ would probably be the actual experimentally measured one: $\chi_0^{exp} \equiv \frac{\partial \langle M \rangle}{\partial H} = g\mu_b \chi_0$. Then there’s the theoretical susceptibility. It’s the susceptibility of the order parameter to the relevant force, which is whatever multiplies the order parameter in the term in the free energy that looks like: $-s(t) \cdot f(t)$. In this problem we take the order parameter to be the spin $S(t)$ and so the relevant force is $f(t) \equiv -g\mu_b H(t)$. The appropriate susceptibility is therefore $\chi_0^{th} \equiv \frac{\partial \langle S \rangle}{\partial f}$. Unfortunately, we also have to distinguish between the static and time-dependent cases. In the time-dependent case the relevant susceptibility is defined by:

$$S(t) = \int_{-\infty}^t \chi^{th}(t-t') f(t') dt'$$

and differs in units from the static susceptibility χ_0^{th} by a factor of 1/time. This is the one that will appear in our fluctuation-dissipation theorem. From this equation we can do a check on units: $\chi^{th}(t)$ has units of 1/(energy · time), since f has units of energy and S is unitless. The appropriate correlation function to match this susceptibility is $C(t) \equiv \langle S(t) S(0) \rangle$, which is also unitless. Hence, the units work out in $\chi^{th} = -\beta \frac{dC(t)}{dt}$. Of course, there is another correlation function that could be of interest—namely $C_{mag}(t) \equiv \langle M(t) M(0) \rangle = (g\mu_b)^2 C(t)$. If using this one, then the appropriate force is no longer $g\mu_b H(t)$ but just $H(t)$, and the order parameter is $M(t)$. In this case, the relevant susceptibility would be $\chi_0^{mag} \equiv \frac{\partial \langle M \rangle}{\partial H} = (g\mu_b)^2 \chi_0^{th}$, and the units still work out, as they must.

(a) In the case ...

We consider a single spin $s = \pm 1/2$ coupled to a heat bath in a static field H , with energy $g\mu_b s H$. Then:

$$\langle S \rangle = -\frac{1}{2} \tanh\left(\frac{g\mu_b \beta H}{2}\right)$$

and:

$$\chi_0 \equiv \frac{\partial \langle S \rangle}{\partial H} = -\frac{1}{4} g\mu_b \beta \operatorname{sech}^2\left(\frac{g\mu_b \beta H}{2}\right) \approx -\frac{g\mu_b \beta}{4} \quad \text{for small } \beta H.$$

(b) Use the Onsager...

Onsager's regression hypothesis says that $C(t)$ obeys the same equation as $\langle S(t) \rangle$:

$$\frac{dC(t)}{dt} = -\Gamma C(t) \quad (H = 0)$$

$$\Rightarrow C(t) = C(0)e^{-\Gamma t} \quad (t \geq 0).$$

Now $C(0) = \langle S(0)^2 \rangle = 1/4$ and $C(t) = C(-t)$, so:

$$C(t) = (1/4)e^{-\Gamma|t|}.$$

It is straightforward to calculate the Fourier transform, assuming $\Gamma \in \mathbb{R}^+$:

$$\tilde{C}(\omega) = \int_{-\infty}^{\infty} dt e^{-i\omega t} C(t) = \dots = \frac{1}{2} \frac{\Gamma}{\Gamma^2 + \omega^2}.$$

(c) Assuming ...

As discussed above, the appropriate susceptibility here is χ^{th} , corresponding to a force $-g\mu_b H(t)$. This is also a straightforward calculation:

$$\chi^{th}(t) = \begin{cases} -\beta \frac{dC(t)}{dt} & t \geq 0 \\ 0 & t < 0 \end{cases} = \dots = \begin{cases} (1/4)\beta\Gamma e^{-\Gamma t} & t \geq 0 \\ 0 & t < 0 \end{cases}$$

And the Fourier transform is found to be:

$$\tilde{\chi}^{th}(\omega) = \dots = \frac{1}{4} \frac{\beta\Gamma}{\Gamma + i\omega}.$$

(d) Compute ...

Power equals force times velocity. Averaged over a cycle, the power dissipated by the spin as a function of the frequency of the external field $H(t)$ is:

$$p(\omega) = \frac{1}{T} \int_0^T \left(-g\mu_b H(t) \frac{dS(t)}{dt} \right) dt.$$

We can find the 'velocity' $\frac{dS(t)}{dt}$ from:

$$S(t) = \int_{-\infty}^t \chi(t-t') (-g\mu_b H(t')) dt'.$$

Plugging in $H(t) = H_0 \cos \omega t$ this becomes:

$$S(t) = -g\mu_b H_0 \left(\text{Re}\{\tilde{\chi}^{th}(\omega)\} \cos \omega t - \text{Im}\{\tilde{\chi}^{th}(\omega)\} \sin \omega t \right).$$

Next we use our formula for $\tilde{\chi}^{th}(\omega)$ from above and differentiate the resulting expression with respect to time. Evaluating the power integral then gives the result:

$$p(\omega) = \frac{\beta \Gamma (g\mu_b H_0 \omega)^2}{8(\Gamma^2 + \omega^2)}.$$

10.5 Telegraph Noise in Nanojunctions.

(a) **Master Equation.** What is the 2×2 matrix \mathbf{M} for our system? At long times, what fraction of the time will our system spend in the α state, $\langle \rho_\alpha \rangle = \lim_{t \rightarrow \infty} \rho_\alpha(t)$?

We write the Master equation $d\vec{\rho}/dt = \mathbf{M} \cdot \vec{\rho}$, where

$$\vec{\rho} = \begin{pmatrix} \rho_\alpha \\ \rho_\beta \end{pmatrix} \quad \text{and} \quad \mathbf{M} = \begin{pmatrix} -P_{\beta\alpha} & P_{\alpha\beta} \\ P_{\beta\alpha} & -P_{\alpha\beta} \end{pmatrix},$$

since $d\rho_\alpha/dt + d\rho_\beta/dt = 0$. At long times, $d\vec{\rho}/dt = 0$ so $P_{\beta\alpha} \langle \rho_\alpha \rangle = P_{\alpha\beta} \langle \rho_\beta \rangle$. We have $\rho_\alpha + \rho_\beta = 1$ so we can solve

$$\langle \rho_\alpha \rangle = \frac{P_{\alpha\beta}}{P_{\alpha\beta} + P_{\beta\alpha}}.$$

(b) Find the eigenvalue-eigenvector pairs for \mathbf{M} . Which corresponds to the stationary state $\vec{\rho}(\infty)$ from part (a)? Suppose that at $t = 0$ the system is known to be in the α state, $\vec{\rho}(0) = (1, 0)$. Write this initial condition in the basis of eigenvectors, and hence give a formula for the subsequent time evolution $\rho_\alpha(t)$. What is the rate of decay to the stationary state?

We diagonalize \mathbf{M} to find the normalized right eigenvectors

$$\rho^* = \frac{1}{P_{\alpha\beta} + P_{\beta\alpha}} \begin{pmatrix} P_{\alpha\beta} \\ P_{\beta\alpha} \end{pmatrix} \quad \lambda^* = 0$$

$$\tilde{\rho} = \frac{1}{P_{\alpha\beta} + P_{\beta\alpha}} \begin{pmatrix} P_{\beta\alpha} \\ -P_{\beta\alpha} \end{pmatrix} \quad \tilde{\lambda} = -(P_{\alpha\beta} + P_{\beta\alpha}).$$

We see that ρ^* is the stationary state while $\tilde{\rho}$ is transient. We write the second eigenstate with a funny normalization for reasons which will become clear now. These eigenstates are complete, and so we can write any state as a linear combination of them. Specifically, the pure α state, $\vec{\rho}(0) = \rho^* + \tilde{\rho}$. Now the time evolution in the eigenbasis is simply $e^{-\lambda}$ for each state, so that the stationary state remains while the transient state decays,

$$\vec{\rho}(t) = \rho^* + e^{-(P_{\alpha\beta} + P_{\beta\alpha})t} \tilde{\rho}.$$

We define $P_{\alpha\alpha}(t)$ to be the probability of being in the state α a time t after it started in α . Therefore we read off the α component of each term,

$$P_{\alpha\alpha}(t) = \rho_\alpha(t) = \frac{1}{P_{\alpha\beta} + P_{\beta\alpha}} \left(P_{\alpha\beta} + P_{\beta\alpha} e^{-(P_{\alpha\beta} + P_{\beta\alpha})t} \right).$$

We see that the limiting behavior is correct and we have a decay into the stationary state with half-life $t_{1/2} = \log 2 / (P_{\alpha\beta} + P_{\beta\alpha})$.

(c) **Correlation Function.** Let $R(t)$ be the resistance as a function of time, hopping between R_α and R_β , as shown in figure 10.15, and let \bar{R} be the time average of the resistance. Write a formula for the connected correlation function $C(t) = \langle (R(t') - \bar{R})(R(t' + t) - \bar{R}) \rangle$ in terms of $P_{\alpha\alpha}(t)$?

Following the hint, let us first look at $C_\beta(t) \equiv \langle (R(t') - R_\beta)(R(t' + t) - R_\beta) \rangle$. These are correlations of fluctuations in the equilibrium ensemble: the initial condition is equilibrium, and the average is over time evolutions of the system including thermal fluctuations. $R(t)$ hops between R_α and R_β , so for given t and t' the only contribution to the average in $C_\beta(t)$ will be when $R(t') = R(t) = R_\alpha$. Thus $C_\beta(t) = (R_\alpha - R_\beta)^2 \rho_\alpha(t') P_{\alpha\alpha}(t)$ and since we begin in equilibrium, $\rho_\alpha(t') = \langle \rho_\alpha \rangle$. So

$$\begin{aligned} C(t) &= \langle (R(t') - \bar{R})(R(t' + t) - \bar{R}) \rangle \\ &= \langle [R(t') - R_\beta - (\bar{R} - R_\beta)][R(t' + t) - R_\beta - (\bar{R} - R_\beta)] \rangle \\ &= (R_\alpha - R_\beta)^2 \langle \rho_\alpha \rangle P_{\alpha\alpha}(t) - (R_\beta - \bar{R})^2 \\ &= (R_\alpha - R_\beta)^2 \langle \rho_\alpha \rangle (P_{\alpha\alpha}(t) - \langle \rho_\alpha \rangle), \end{aligned}$$

where we have used the fact that $\langle R(t') \rangle = \bar{R} = R_\alpha \langle \rho_\alpha \rangle + R_\beta \langle \rho_\beta \rangle$ and $(R_\beta - \bar{R})^2 = (R_\alpha - R_\beta)^2 \langle \rho_\alpha \rangle^2$. This satisfies our limiting conditions, $C(\infty) = 0$ and $C(0) = \langle \rho_\alpha \rangle (R_\alpha - \bar{R})^2 + \langle \rho_\beta \rangle (R_\beta - \bar{R})^2$.

(d) **Detailed Balance.** What is the ratio between the two rates $P_{\alpha\beta}$ and $P_{\beta\alpha}$, assuming that the system satisfies detailed balance, in terms of these four rates?

For a three-state system, detailed balance asserts that the rates around the two 'circular' cycle directions are equal: $P_{\alpha\beta}P_{\beta\gamma}P_{\gamma\alpha} = P_{\alpha\gamma}P_{\gamma\beta}P_{\beta\alpha}$. Therefore,

$$\frac{P_{\alpha\beta}}{P_{\beta\alpha}} = \frac{P_{\alpha\gamma}P_{\gamma\beta}}{P_{\beta\gamma}P_{\gamma\alpha}}. \quad (*)$$

While the rates $P_{\alpha\beta}$ and $P_{\beta\alpha}$ are small, they are presumably non-zero. The point of this is that if the system is in equilibrium and satisfies detailed balance, then the ratios of the rates must satisfy (*). Whether they do for a real experimental system is therefore a check on whether the system is in equilibrium.

10.6 Fluctuations-Dissipation: Ising.

(a) Does the shape and the time-scale of the magnetization decay look the same as that of the autocorrelation function? Note down the values for $c(0)$, $C(0)$, $m(0)$, and $M(0)$.

The shape and time-scale of the magnetization decay should be very similar to those of the autocorrelation. Example of numbers: $c(0) = C(0)/(\# \text{ spins})^2 \approx 0.0003$. Careful: the number of spins for a 200×200 lattice is 200^2 , so $C(0) = c(0) \times (200)^4 \approx 4.8 \times 10^5$. Similarly, $m(0) \approx 0.19 \Rightarrow M(0) = m(0) \times (200)^2 \approx 7.6 \times 10^3$.

(b) **Response Functions and the Fluctuation-Dissipation Theorem.** Using equation 10.93, write $M(t)$ for the step down $H(t) = H_0\Theta(-t)$, in terms of $\chi(t)$.

For the step down $H(t) = H_0 \Theta(-t)$ (and with $\langle M \rangle = 0$),

$$M(t) = \int_{-\infty}^t dt' \chi(t-t') H(t') = H_0 \int_t^{\infty} dt' \chi(t').$$

(c) Use equation 10.94 and your answer to part (b) to predict the relationship between the demagnetization $M(t)$ and the correlation $C(t)$ you measured in part (a). How does your analytical ratio compare with the $t = 0$ ratio you noted down in part (a)?

If we plug the fluctuation-dissipation theorem into the previous result we find

$$M(t) = -\beta H_0 \int_t^{\infty} dt' \frac{dC(t')}{dt'} = \beta H_0 (C(t) - C(\infty)) = \beta H_0 G(t)$$

since $C(\infty) = 0$. For $\beta = 1/3$ and $H_0 = 0.05$ we therefore expect $M(0)/C(0) \approx 0.017$, which is indeed close to what we find with the sample numbers from part (a), $7.6 \times 10^3 / 4.8 \times 10^5 \approx 0.016$.

10.7 Noise and Langevin equations.

(a) Integrating by parts and keeping the boundary terms, show that the equation of motion has the form

$$m\ddot{Q} = -\frac{\partial \bar{V}}{\partial Q} + \mathbb{F}_b - \beta \int_{-\infty}^t dt' C_b(t-t') \dot{Q}(t').$$

What is the ‘potential of mean force’ \bar{V} , in terms of V and C_b ?

We begin with equations of motion

$$m\ddot{Q} = -\frac{\partial V}{\partial Q} + \mathbb{F}_b + \int_{-\infty}^t dt' \chi_b(t-t') Q(t').$$

We integrate by parts, setting $u = Q(t')$ and $dv = \chi_b(t-t') dt'$, so that $du = \dot{Q}(t') dt'$ and by the fluctuation-dissipation theorem, $v = \beta C_b(t-t')$. Thus,

$$\int_{-\infty}^t u dv = \beta Q(t') C_b(t-t') \Big|_{-\infty}^t - \int_{-\infty}^t dt' \beta C_b(t-t') \dot{Q}(t') = \beta Q(t) C_b(0) - \int_{-\infty}^t dt' \beta C_b(t-t') \dot{Q}(t'),$$

since $C_b(\infty) \rightarrow 0$. If we define $\bar{V} = V - \frac{1}{2} \beta C_b(0) Q(t)^2$ then the equations of motion become

$$m\ddot{Q} = -\frac{\partial \bar{V}}{\partial Q} + \mathbb{F}_b - \int_{-\infty}^t dt' \beta C_b(t-t') \dot{Q}(t'),$$

and we see that \bar{V} is the ‘potential of mean force’.

(b) If the correlations in the bath are short-lived compared to the time-scales of the system, we can approximate $\dot{Q}(t') \approx \dot{Q}(t)$ in equation 10.90, leading to a viscous friction force $-\gamma \dot{Q}$. What is the formula for γ ?

We now take $\dot{Q}(t') = \dot{Q}(t)$. Then $\dot{Q}(t)$ factors out of the integral, leaving

$$m\ddot{Q} = -\frac{\partial \bar{V}}{\partial Q} + \mathbb{F}_b - \left(\int_{-\infty}^t dt' \beta C_b(t-t') \right) \dot{Q}(t) = -\frac{\partial \bar{V}}{\partial Q} + \mathbb{F}_b - \gamma(t) \dot{Q}(t),$$

where we define the viscous friction as

$$\gamma = \int_{-\infty}^t dt' \beta C_b(t-t') = \int_0^{\infty} dt' \beta C_b(t').$$

(c) Conversely, for a model system with a perfect viscous friction law $-\gamma\dot{Q}$ at temperature T , derive the equation for the correlation function for the noise $C_b(t-t')$.

If instead of the Langevin term we include a term $-\gamma\dot{Q}(t)$, we can identify this equation with our result from (a), (10.90) in the text. Then $\gamma\dot{Q}(t) = \int_{-\infty}^t dt' \beta C_b(t-t')\dot{Q}(t')$. This can only hold if $\beta C_b(t-t') = \delta(t-t')$ and therefore we find that for a perfect viscous friction law,

$$C_b(t) = \frac{\gamma}{\beta} \delta(t).$$

Thus, the bath has no memory and is Markovian.

10.8 Magnet Dynamics.

(a) Show that the free energy $F[M]$, when written in terms of \tilde{M}_n , becomes an independent sum over modes, with a quadratic energy in each mode. What is $\langle |\tilde{M}_n|^2 \rangle_{\text{eq}}$, by equipartition? Argue that

$$\langle \tilde{M}_{-m} \tilde{M}_n \rangle_{\text{eq}} = \frac{k_B T}{L(Ck_n^2 + B)} \delta_{mn}.$$

(Thanks to Duane Loh for assistance with this solution.)

Keeping track of the independent degrees of freedom in complex notation is subtle. (Or rather, it works simply for subtle reasons.) So, we'll instead write $M(x)$ and \mathcal{F} in terms of real Fourier modes (cosine and sine transforms), use the equipartition theorem to estimate their amplitudes, and then find the equilibrium value for $\langle \tilde{M}_n \tilde{M}_n^* \rangle_{\text{eq}}$. We'll then see that this equilibrium formula looks much the same as one might have naively calculated in complex notation, ignoring the subtleties.

We begin with our naive complex representation of the magnetic order parameter:

$$M(x) = \sum_{n=-N/2}^{+N/2} \tilde{M}_n e^{ik_n x},$$

where $k_n = \frac{2\pi n}{L}$ with $L/N = a$, the lattice constant. If N were even, then the limits of summation would run from $N/2 - 1$ to $N/2$. We assume for simplicity that N is odd for a more symmetric notation. Since $M(x)$ is real, representing

it as a sum of complex exponentials requires Friedel symmetry (i.e. $\widetilde{M}_n = \widetilde{M}_{-n}^*$):

$$\begin{aligned} \text{For } n = 0, \quad & \widetilde{M}_0 = u_0 \quad \text{and we require } v_0 = 0 \\ \text{For } n > 0, \quad & \widetilde{M}_n = u_n - iv_n \\ \text{For } n < 0, \quad & \widetilde{M}_n = u_{-n} + iv_{-n}, \end{aligned}$$

where the only independent degrees of freedom are the real variables: u_n and v_n for $n \geq 0$. Then expanding $M(x)$ in terms of sines and cosines:

$$M(x) = u_0 + \sum_{n>0} 2 [u_n \cos(k_n x) + v_n \sin(k_n x)] \quad (2)$$

where the factor of 2 comes for combining the $\sum_{n>0}$ terms with the related $\sum_{n<0}$ terms. From our expression for \mathcal{F} :

$$\begin{aligned} \mathcal{F} = & \frac{C}{2} \left[\sum_{n>0} 2k_n (-u_n \sin(k_n x) + v_n \cos(k_n x)) \right]^2 \\ & + \frac{B}{2} \left[\sum_{n>0} 2(u_n \cos(k_n x) + v_n \sin(k_n x)) \right]^2. \end{aligned}$$

Using the orthogonality of sines and cosines

$$\int_0^L \sin(k_n x) \sin(k_m x) dx = \frac{L}{2} \delta_{nm} = \int_0^L \cos(k_n x) \cos(k_m x) dx \quad (3)$$

and zero for all other combinations sines and cosines, we find that

$$F = \int \mathcal{F} dx = Lu_0^2 \frac{B}{2} + L \sum_{n>0} [u_n^2 + v_n^2] (Ck_n^2 + B) \quad (4)$$

Taking the expectation values over all equilibrium states:

$$\langle F \rangle_{eq} = L \langle u_0^2 \rangle_{eq} \frac{B}{2} + L \sum_{n>0} [\langle u_n^2 \rangle_{eq} + \langle v_n^2 \rangle_{eq}] (Ck_n^2 + B) \quad (5)$$

Classically, if all the modes are excited at equilibrium, by the equi-partition theorem:

$$\langle u_0^2 \rangle_{eq} = \frac{k_B T}{LB}, \quad \langle u_{n>0}^2 \rangle_{eq} = \langle v_{n>0}^2 \rangle_{eq} = \frac{k_B T}{2L(Ck_n^2 + B)}. \quad (6)$$

Now we examine the degrees of freedom in $\langle \widetilde{M}_n \widetilde{M}_m^* \rangle_{eq}$ for all possible cases of n and m . We first recall that $\langle \widetilde{M}_n \widetilde{M}_m^* \rangle_{eq} = \langle \widetilde{M}_n \widetilde{M}_{-m} \rangle_{eq}$ by Friedel symmetry. Clearly, $\langle \widetilde{M}_n \widetilde{M}_m^* \rangle_{eq} = 0$ unless $n = \pm m$ because u 's and v 's are independent ($\widetilde{M}_{n>0} = u_{n>0} - iv_{n>0}$ and $\widetilde{M}_{n<0} = u_{-n>0} + iv_{-n>0}$).

$$\begin{aligned} n = -m > 0, \quad & \langle \widetilde{M}_n \widetilde{M}_m^* \rangle_{eq} = \langle u_n - iv_n \rangle_{eq} \langle (u_{-m} + iv_{-m})^* \rangle_{eq} \\ & = \langle u_n^2 \rangle_{eq} - \langle v_n^2 \rangle_{eq} - 2i \langle u_n v_n \rangle = 0. \end{aligned}$$

Similarly, the cases where $-n = m > 0$ vanishes too. Next we consider:

$$\begin{aligned} n = m > 0, \quad & \langle \widetilde{M}_n \widetilde{M}_m^* \rangle_{eq} = \langle u_n - iv_n \rangle_{eq} \langle (u_m - iv_m)^* \rangle_{eq} \\ & = \langle u_n^2 \rangle_{eq} + \langle v_n^2 \rangle_{eq} - 2i \langle u_n v_n \rangle = 0 \end{aligned}$$

Again, we get the same result for $n = m < 0$. But for $n = m = 0$, $v_n = 0$ for $M(x)$ to be real. Hence:

$$\langle \widetilde{M}_0 \widetilde{M}_0^* \rangle_{eq} = \langle u_0^2 \rangle_{eq} \quad (7)$$

Combining what we have found so far:

$$\langle \widetilde{M}_n \widetilde{M}_m^* \rangle_{eq} = \boxed{\langle \widetilde{M}_n \widetilde{M}_{-m} \rangle_{eq} = \frac{k_B T}{L(Ck_n^2 + B)} \delta_{nm}}. \quad (8)$$

Alternatively, if we treated \widetilde{M}_n as a single degree of freedom, naively ignoring Friedel symmetry and the fact that \widetilde{M}_n is complex, applying equipartition theorem, we arrive at the same answer for $\langle \widetilde{M}_n \widetilde{M}_m^* \rangle_{eq}$.

$$\begin{aligned} \mathcal{F} &= \frac{C}{2} (\nabla M)^2 + \frac{B}{2} (M)^2 \\ &= \frac{C}{2} (\nabla M) (\nabla M)^* + \frac{B}{2} (M) (M)^* \\ &= \sum_n \sum_m \widetilde{M}_n \widetilde{M}_m^* e^{i(k_m - k_n)x} \left(\frac{C}{2} k_n k_m + \frac{B}{2} \right) \end{aligned} \quad (9)$$

$$F = \int \mathcal{F} dx = L \sum_n \widetilde{M}_n \widetilde{M}_n^* \left(\frac{C}{2} k_n^2 + \frac{B}{2} \right) \quad (10)$$

If we ignore the Friedel symmetry $\widetilde{M}_n^* = \widetilde{M}_{-n}$, and treat \widetilde{M}_n as if it had only one degree of freedom, then the equipartition theorem would suggest that

$$\frac{k_B T}{2} = L \left(\frac{C}{2} k_n^2 + \frac{B}{2} \right) \langle \widetilde{M}_n \widetilde{M}_n^* \rangle_{eq}, \quad (11)$$

giving us the right answer due to a subtle cancellation of errors. Ignoring the 2 degrees of freedom in \widetilde{M}_n , but assuming that Friedel pairs are independent ($k_{\pm n}$ modes are independent), is the traditional shortcut which gives us the correct equilibrium result.

(b) Calculate the equal-time correlation function for the magnetization in equilibrium, $C(x, 0) = \langle M(x, 0) M(0, 0) \rangle_{eq}$.

Assuming the interchange of the averaging and infinite sums is ok,

$$C(x, 0) = \langle M(x, 0) M(0, 0) \rangle = \sum_{n, m=-\infty}^{\infty} \langle \widetilde{M}_n \widetilde{M}_m \rangle e^{ik_n x} = \frac{k_B T}{L} \sum_n \frac{e^{ik_n x}}{Ck_n^2 + B} = \frac{k_B T}{2\pi} \int_{-\infty}^{\infty} dk \frac{e^{ikx}}{Ck^2 + B}.$$

In changing the sum to an integral, we are in effect taking the limit $L \rightarrow \infty$. Using the formula for the Fourier transform given in the text, we find

$$C(x, 0) = \frac{k_B T}{2\sqrt{BC}} e^{-\sqrt{B/C}|x|}.$$

(c) Calculate $\delta \mathcal{F} / \delta M$.

We use the Euler-Lagrange equation,

$$\delta \mathcal{F} = \frac{\delta \mathcal{F}}{\delta M} \delta M = \left(\frac{\partial \mathcal{F}}{\partial M} - \nabla \cdot \frac{\partial \mathcal{F}}{\partial \nabla M} \right) \delta M$$

so that

$$\frac{\delta \mathcal{F}}{\delta M} = BM - C\nabla^2 M.$$

Notice that spins do not typically move in the direction of an applied force. In many systems, spins precess under external forces, moving in a direction roughly perpendicular to the applied force and only slowly drifting downhill under the effects of dissipative. We are modeling an overdamped system, where this precession is slow compared to the damping.

(d) From your answer to part (c), calculate the Greens function for $[M]_{M_i}$, $G(x, t)$ giving the time evolution of an initial condition $M_i(x) = M(x, 0) = G(x, 0) = \delta(x)$.

The Green's function is the solution to the equation of motion for a delta function initial condition, $M(x, 0) = G(x, 0) = \delta(x)$. Since the equation is linear and taking derivatives commutes with averaging, it applies equally well to $M(x, t)$ as to $\langle M(x, t) \rangle$,

$$\frac{\partial M}{\partial t} = -\eta \frac{\delta \mathcal{F}}{\delta M} = \eta(C\nabla^2 M - BM).$$

We can solve this with Fourier transforms. First, substitute the Fourier transform of M ,

$$M(x, t) = \frac{1}{2\pi} \int dk e^{ikx} \tilde{M}(k, t).$$

This gives us

$$\int dk e^{ikx} \left[\frac{\partial \tilde{M}}{\partial t} + \eta(Ck^2 + B)\tilde{M} \right] = 0$$

$$\frac{\partial \tilde{M}}{\partial t} = -\eta(B + Ck^2)\tilde{M}$$

$$\tilde{M}(k, t) = \tilde{M}(k, 0)e^{-\eta(B+Ck^2)t}.$$

Now we specialize to the Green's function — i.e. the initial condition corresponding to a delta function, $\tilde{G}(k, 0) = \tilde{\delta}(k) = 1$. Applying the inverse Fourier transform,

$$G(x, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ikx} e^{-\eta(B+Ck^2)t}.$$

This integral can be solved by completing the square in the exponent and then changing variables until one arrives at

$$G(x, t) = \frac{1}{2\pi\sqrt{\eta Ct}} e^{-\eta Bt - \frac{x^2}{4\eta Ct}} \int_{-\infty - ix/2\sqrt{\eta Ct}}^{\infty - ix/2\sqrt{\eta Ct}} d\nu e^{-\nu^2}.$$

It is a standard trick in complex analysis to show that

$$\int_{-\infty - iA}^{\infty - iA} dy e^{-y^2} = \int_{-\infty}^{\infty} dy e^{-y^2} = \sqrt{\pi},$$

because the contribution to the integral from the two ends of the rectangular contour is zero. Thus we obtain the Green's function:

$$G(x, t) = \frac{1}{2\sqrt{\eta\pi Ct}} e^{-\left(\eta Bt + \frac{x^2}{4\eta Ct}\right)}.$$

(e) Using the Onsager regression hypothesis calculate the space-time correlation function $C(x, t) = \langle M(x, t)M(0, 0) \rangle_{\text{ev}}$. If it's convenient, plot it for short times and for long times: does it look like $\exp(-|y|)$ in one limit and $\exp(-y^2)$ in another?

Onsager's regression hypothesis says that fluctuations in the equilibrium ensemble (correlation function) evolve in the same way as externally imposed perturbations. Thus, the Green's function for $M(x, t)$ is also the Green's function for $C(x, t)$,

$$C(x, t) = \int_{-\infty}^{\infty} C(x', 0)G(x - x', t)dx' = \frac{k_B T}{4C\sqrt{\eta\pi Bt}} \int_{-\infty}^{\infty} e^{-\sqrt{\frac{B}{C}}|x'|} e^{-(\eta Bt + \frac{(x-x')^2}{4\eta C t})} dx'.$$

This is a moderately long integral to do by hand (or you can plug it into Mathematica or a similar program). Using the error function defined as

$$\text{erf } t \equiv \frac{2}{\sqrt{\pi}} \int_0^t dz e^{-z^2},$$

we eventually arrive at the result,

$$C(x, t) = \frac{k_B T}{4\sqrt{BC}} \left\{ e^{-\sqrt{\frac{B}{C}}x} \left[1 + \text{erf} \left(\frac{x}{2\sqrt{\eta C t}} - \sqrt{\eta B t} \right) \right] + e^{\sqrt{\frac{B}{C}}x} \left[1 - \text{erf} \left(\frac{x}{2\sqrt{\eta C t}} + \sqrt{\eta B t} \right) \right] \right\}.$$

(f) Calculate the susceptibility $\chi(x, t)$ from $C(x, t)$. Start by giving the abstract formula, and then plug in your answer from part (e).

To calculate the susceptibility from the correlation function we simply use

$$\chi(x, t) = -\beta \frac{\partial C(x, t)}{\partial t} \quad (t \geq 0)$$

and remember that $\chi = 0$ for $t < 0$. Along the way we will also need

$$\frac{d}{dt} \text{erf}(f(t)) = \frac{2}{\sqrt{\pi}} \frac{df}{dt} e^{-f^2}.$$

This is a straightforward calculation to do and it happens in this case that

$$\chi(x, t) = \begin{cases} \eta G(x, t) & t \geq 0 \\ 0 & t < 0 \end{cases}$$

10.9 Quasiparticle poles and Goldstone's theorem.

(a) Solve for Ω_k and find the new dispersion relation. With what power of k does the quality factor diverge at small k ?

By substituting the ansatz in the equation, we find:

$$\Omega_k = \omega_k - i\Gamma_k$$

$$\Gamma_k = \frac{k^2 d^2}{2}$$

$$\omega_k = k \sqrt{c^2 - \left(\frac{k d^2}{2}\right)^2}$$

At long wavelengths, it is clear that $\omega_k \rightarrow \pm ck$ and the quality factor diverges as k^{-1} .

(b) At what wavelength does the real part of the frequency vanish?

For $k = \frac{2c}{d^2}$, $\omega_k = 0$.

(c) What is the susceptibility of the damped system?

using $u(k, \omega) = \chi(k, \omega) f(k, \omega)$ we get:

$$\chi(k, \omega) = \frac{1}{\rho} (c^2 k^2 - \omega^2 - i d^2 \omega k^2)^{-1}$$

When $d = 0$, the poles are at $\pm ck$. When d no longer vanishes, the poles move to $\pm \omega_k - i\Gamma_k$.

(d) Check that the poles in the susceptibility are all in the lower half-plane, as required by causality.

The imaginary part of the poles is $-\Gamma_k$, which is smaller than 0. The poles are thus in the lower half plane.

(e) Calculate $C(k, \omega)$ from the fluctuation-dissipation theorem.

$$C(k, \omega) = \frac{2k_B T k^2 d^2}{\rho (c^2 k^2 - \omega^2)^2 + (d^2 k^2 \omega)^2}$$

The dispersion relation is the function ω_k , which should be easy to deduce from $C(k, \omega)$ (for the parameters given), since C has such a sharp peak at that frequency.

11.1 Maxwell and Van Der Waals.

Lines on a graph...

11.4 Nucleation in the Ising Model.

(a) Using the right-hand graph of magnetization versus time, measure the average time it takes to cross zero by averaging ten measurements. Similarly measure the average time to nucleate the down phase for ($H_{\text{ext}} = -0.2$), changing number of sweeps to 10^4 and speed to 100. Since the nucleation center can be located at any site on the lattice, the nucleation rate scales with the number of spins in the system. Calculate, for both fields, the nucleation rate per spin $\Gamma_{\text{exp}}(H)$.

The nucleation rate per spin $\Gamma_{\text{exp}}(H_0) \sim 1/t_{\text{avg}}$, where t_{avg} is the average time for the magnetization to cross zero. Some typical numbers for this part:

$$\Gamma_{\text{exp}}(H = -0.3) \approx 1.7 \times 10^{-6} \frac{1}{\text{spin} \cdot \text{sweep}}, \quad \Gamma_{\text{exp}}(H = -0.2) \approx 1 \times 10^{-7} \frac{1}{\text{spin} \cdot \text{sweep}}.$$

(b) Give the formula for the free energy of a flipped cluster of radius R as a function of σ , H , and ΔM . Give formulas for R_c (the critical droplet size where the free energy is a local maximum), the resulting barrier B to nucleation, and the predicted rate $\Gamma_{\text{theory}} = \exp(-B/T)$ (assuming a prefactor of roughly one attempt per sweep per spin). At low temperatures, $\sigma \sim 2J \equiv 2$ and $\Delta M \approx 2$, since the system is almost fully magnetized and σ is the number of broken bonds ($2J$ each) per unit length of interface. Make a table with rows for the two fields you simulated and with columns for H , R_c , B , Γ_{theory} , and Γ_{exp} from (a).

The free energy of a flipped cluster of radius R is:

$$F = 2\pi R\sigma - H_{\text{ext}}\Delta M\pi R^2.$$

Note that we are working in two dimensions, so the 'surface tension' σ has units of force instead of the usual force per length in 3D. We find the critical radius for nucleation:

$$\frac{dF}{dR} = 0 \quad \Longleftrightarrow \quad R_c = \frac{\sigma}{H_{\text{ext}}\Delta M}.$$

The resulting barrier to nucleation is the free energy at the critical radius:

$$B = F(R_c) = \frac{\pi\sigma^2}{H_{\text{ext}}\Delta M}.$$

From this we find the theoretical rate per spin per sweep:

$$\Gamma_{\text{theory}} \sim \exp\left(-\frac{\pi\sigma^2}{H_{\text{ext}}\Delta MT}\right).$$

With $\sigma \approx 2$ and $\Delta M \approx 2$ most people didn't get good agreement: in most cases the Γ differed by a factor of 10 or 100.

(c) Start with $H = -0.2$, $T = 1.5$ and a down-spin droplet of radius five (diameter of ten). Does it grow more often than it shrinks, or vice-versa? Make a table of the values of H and f you measure. Vary the field H until the probabilities roughly match: find the field for $R_c = 5$ to within 0.1. For what field is the theoretical critical droplet radius $R_c = 5$ at $T = 1.5$?

The theoretical field giving a critical radius of 5 is $H_{\text{theory}} = \sigma/(\Delta M R_c) \approx 2/(2 \cdot 5) = .2$, while most found $H_{\text{exp}} \approx .15$, which is close.

11.5 Nucleation of Dislocation Pairs.

Of the following statements, which are true?

(T) (F) The critical radius R_c is proportional to $1/\sigma_{xy}$.

True. The effective free energy for the dislocation pair is $F = C + \frac{\mu a^2}{2\pi(1-\nu)} \log\left(\frac{R}{a}\right) - \sigma_{xy} R a$, where μ , ν , and a are constants that depend on the material, and C is the fixed core energy for the defect pair. In analogy with critical droplet problems, we can find the critical radius for nucleation by differentiating $dF/dR = 0$ so that

$$R_c = \frac{\mu a}{2\pi(1-\nu)\sigma_{xy}} \propto \frac{1}{\sigma_{xy}}.$$

(T) (F) The energy barrier to thermal nucleation is proportional to $1/\sigma_{xy}^2$.

False. The energy barrier is then given by

$$B = F(R_c) = (\text{Constant}) + \frac{\mu a^2}{2\pi(1-\nu)} \log\left(\frac{\mu}{\sigma_{xy}}\right) \not\propto \frac{1}{\sigma_{xy}^2}.$$

(T) (F) The rate Γ of thermal nucleation of dislocations predicted by our critical droplet calculation is of the form $\Gamma = \Gamma_0(T)(\sigma_{xy}/\mu)^{D/k_B T}$, for a suitable material-dependent function $\Gamma_0(T)$ and constant D .

True. The resulting rate of thermal nucleation of dislocation pairs is

$$\Gamma \propto \exp\left(-\frac{B}{k_B T}\right) \propto \left(\frac{\sigma_{xy}}{\mu}\right)^{D/k_B T},$$

where $D = \mu a^2/2\pi(1-\nu)$ is a material-dependent constant.

(T) (F) According to our calculation, the response of a two-dimensional crystal under stress is indistinguishable from that of a liquid: even at low temperatures, the strain rate due to an external shear force is proportional to the stress.

False. The rate of nucleation of pairs is basically the rate of strain of the crystal – for each defect pair that exceeds R_c in separation the crystal lattice shifts by a . And from the last equation above we see that the strain rate is proportional to some temperature-dependent power of the shear stress σ_{xy} . This is different from a liquid which we're told has $\Gamma \propto \sigma_{xy}$ at low T .

11.5 Nucleation of Dislocation Pairs.

Of the following statements, which are true?

(T) (F) The critical radius R_c is proportional to $1/\sigma_{xy}$.

True. The effective free energy for the dislocation pair is $F = C + \frac{\mu a^2}{2\pi(1-\nu)} \log\left(\frac{R}{a}\right) - \sigma_{xy}Ra$, where μ , ν , and a are constants that depend on the material, and C is the fixed core energy for the defect pair. In analogy with critical droplet problems, we can find the critical radius for nucleation by differentiating $dF/dR = 0$ so that

$$R_c = \frac{\mu a}{2\pi(1-\nu)\sigma_{xy}} \propto \frac{1}{\sigma_{xy}}.$$

(T) (F) The energy barrier to thermal nucleation is proportional to $1/\sigma_{xy}^2$.

False. The energy barrier is then given by

$$B = F(R_c) = (\text{Constant}) + \frac{\mu a^2}{2\pi(1-\nu)} \log\left(\frac{\mu}{\sigma_{xy}}\right) \not\propto \frac{1}{\sigma_{xy}^2}.$$

(T) (F) The rate Γ of thermal nucleation of dislocations predicted by our critical droplet calculation is of the form $\Gamma = \Gamma_0(T)(\sigma_{xy}/\mu)^{D/k_B T}$, for a suitable material-dependent function $\Gamma_0(T)$ and constant D .

True. The resulting rate of thermal nucleation of dislocation pairs is

$$\Gamma \propto \exp\left(-\frac{B}{k_B T}\right) \propto \left(\frac{\sigma_{xy}}{\mu}\right)^{D/k_B T},$$

where $D = \mu a^2/2\pi(1-\nu)$ is a material-dependent constant.

(T) (F) According to our calculation, the response of a two-dimensional crystal under stress is indistinguishable from that of a liquid: even at low temperatures, the strain rate due to an external shear force is proportional to the stress.

False. The strain rate is given roughly by the number of dislocations, which is determined by a balance between the nucleation rate you calculated and the annihilation rate. Since the nucleation rate is a nonlinear power-law function of the

external stress σ_{xy} , the strain rate is also presumably nonlinear – hence different from a liquid which we are told has $\Gamma \propto \sigma_{xy}$ at low T . Note that at high enough temperatures the nucleation rate could be linear, if $D = k_B T$ so $\mu a^2 / k_B T = 2\pi(1-\nu)$. However, this is probably normally above the melting temperature. $\nu \approx 1/2$ for two-dimensional crystals, so the right-hand side is about three; the left-hand side $\mu a^2 / k_B T \approx 15$ at the melting temperature for two dimensional colloidal crystals (A. Wille, F. Valmont, K. Zahn, and G. Maret, Shear modulus of two-dimensional colloidal crystals, *Europhys. Lett.*, **57**, 219 (2002)).

11.6 Coarsening in the Ising Model.

(a) Argue that at zero temperature the total energy above the ground state energy is proportional to the perimeter separating up-spin and down-spin regions. (At finite temperatures, there is a contribution from thermally flipped spins, which shouldn't really count as perimeter for coarsening.) Argue that the inverse of the perimeter per unit area is a reasonable definition for the length-scale of the pattern.

At zero temperature and zero external field, the ground state energy is with all spins aligned either up or down. The energy above ground will be due exclusively to spin-spin interactions for anti-parallel spins. These occur only on the boundary between up-spin and down-spin regions. Therefore, the total energy above ground is proportional to the perimeter separating these regions. Roughly, a very large perimeter would correspond to lots of little droplets of spins, while a small perimeter would correspond to a few large blobs. It is reasonable, therefore, to take the inverse of the perimeter as a measure of the length-scale of the pattern.

(b) Initialize at $T=\text{inf}$, set temperature and external field to zero and number of sweeps to one, run for one sweep, and measure the mean energy per unit area $\langle E \rangle$. Measure the mean energy as a function of time for $t = 2, 4, 8, \dots, 1024$: reset/run/measure $\langle E \rangle$ /double number of sweeps/repeat. (You're measuring the average perimeter length over the last half of the time interval, but that scales in the same way as the perimeter does.) Make a table with columns for t , $\langle E(t) \rangle$, and $L(t) \propto 1/(\langle E \rangle + 2)$. Make a log-log plot of your estimate of the coarsening length $L(t) \propto 1/(\langle E \rangle + 2)$ versus time. What power law does it grow with? What power law did we expect?

One should find $L(t) \sim t^{1/2}$, which is what we expect for coarsening with a non-conserved order parameter.

11.7 Origami Microstructures.

(a) Show that the zero energy ground states of the paper free energy density (equation 11.26) include the two variants and rotations thereof, as shown in figure 11.16. Specifically, show that any rotation $y_i(x_j) = R_{ij}x_j$ of the lined-side-up position is a ground state, where $R_{ij} = \begin{pmatrix} \cos \theta_\ell & -\sin \theta_\ell \\ \sin \theta_\ell & \cos \theta_\ell \end{pmatrix}$ and also that flipping the paper to the white-side-up and then rotating $y_i(x_j) = R_{ik}P_{kj}x_j = \begin{pmatrix} \cos \theta_w & -\sin \theta_w \\ \sin \theta_w & \cos \theta_w \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$ gives a ground state.

Note: The problem was changed considerably for clarity. This solution addressees the nuances of the less clear version, in which P is a flip across the y axis and R rotates clockwise. The new version has P flip across the x axis and R rotates counter clockwise. But the results are the same either way, as long as we are consistent.

We have a free energy density $\mathcal{F} = \alpha(\partial_i y_j \partial_i y_k - \delta_{jk})^2$. Consider first a rotation of the paper right-side up. Then $y_j(x_k) = R_{jk}x_k$ and

$$\partial_i y_j(x_k) = \partial_i R_{jk}x_k = R_{jk} \frac{\partial x_k}{\partial x_i} = R_{jk} \delta_{ki} = R_{ji}.$$

This leaves $\mathcal{F} = \alpha(R_{ji}R_{ki} - \delta_{jk}) = \alpha(R_{ji}R_{ik}^T - \delta_{jk}) = \alpha([RR^T]_{jk} - \delta_{jk})$. The rotation matrix is orthogonal,

$$[RR^T]_{ij} = \begin{pmatrix} \cos \theta_\ell & \sin \theta_\ell \\ -\sin \theta_\ell & \cos \theta_\ell \end{pmatrix} \begin{pmatrix} \cos \theta_\ell & -\sin \theta_\ell \\ \sin \theta_\ell & \cos \theta_\ell \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \delta_{ij}$$

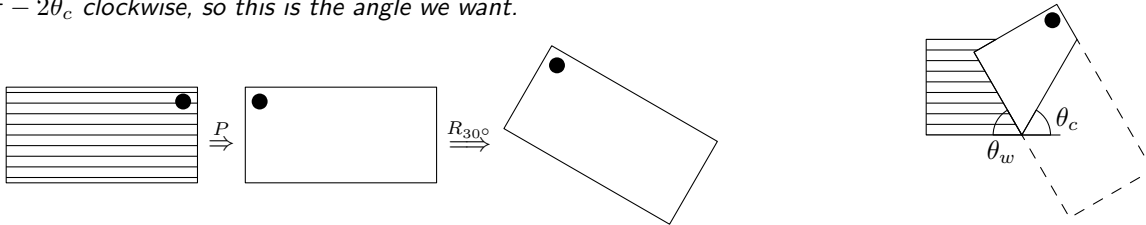
so that $\mathcal{F} = 0$. We do the same calculation for the flipped rotation, $y_i(x_j) = R_{ik}P_{kj}x_j = Q_{ij}x_j$ where

$$Q_{ij} = R_{ik}P_{kj} = \begin{pmatrix} \cos \theta_w & \sin \theta_w \\ -\sin \theta_w & \cos \theta_w \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -\cos \theta_w & \sin \theta_w \\ \sin \theta_w & \cos \theta_w \end{pmatrix}.$$

Again the partial derivative in \mathcal{F} passes through the matrices so that $\mathcal{F} = \alpha([QQ^T]_{ij} - \delta_{ij})$, where we have again $[QQ^T]_{ij} = \delta_{ij}$ is the identity matrix. Again, $\mathcal{F} = 0$. Thus, rotations from both sides of the paper result in the zero-energy ground state.

(b) Place a piece of paper long-edge downward on the table. Holding the left end fixed at $\theta_\ell = 0$, try folding it along crease lines at different angles θ_c . Find a definite relation between the crease angle θ_c and the angle θ_w of the right-hand portion of the paper.

After thorough experimentation we can write $\theta_w = \pi - 2\theta_c$ where θ_c is counter-clockwise from the positive x -axis and θ_w is clockwise from $-x$. We pay close attention now to how exactly these transformations act. Below are two figures, the first displaying the action of the flip P followed by a 30° rotation R_{30° . We note that R rotates clockwise. This is easy to see by applying it to e.g. $(1, 0)$. We see in the second diagram that the crease causes the white side to be rotated by $\theta_w = \pi - 2\theta_c$ clockwise, so this is the angle we want.



(c) Given the relation you deduced for the geometry in part (b), show that the difference in the directional derivatives $(D\mathbf{y}^\ell - D\mathbf{y}^w)$ is zero along \mathbf{c} , $(D\mathbf{y}^\ell - D\mathbf{y}^w) \cdot \mathbf{c} = (\partial_j y_i^\ell - \partial_j y_i^w)c_j = 0$.

We have $y_i^\ell = x_i$ since we're not doing anything to the front. Therefore $\partial_j y_i^\ell = \delta_{ij}$ is the identity. From part (b) we find $\partial_j y_i^w = R_{ik}P_{kj} = Q_{ij}$. Taking $\theta_w = \pi - 2\theta_c$,

$$\delta_{ij} - Q_{ij} = \begin{pmatrix} 1 + \cos(\pi - 2\theta_c) & -\sin(\pi - 2\theta_c) \\ -\sin(\pi - 2\theta_c) & 1 - \cos(\pi - 2\theta_c) \end{pmatrix} = \begin{pmatrix} 1 - \cos 2\theta_c & -\sin 2\theta_c \\ -\sin 2\theta_c & 1 + \cos 2\theta_c \end{pmatrix} = \begin{pmatrix} 2\sin^2 \theta_c & -2\sin \theta_c \cos \theta_c \\ -2\sin \theta_c \cos \theta_c & 2\cos^2 \theta_c \end{pmatrix}$$

where we have used several trig identities along the way. Now $\mathbf{c}_j = (\cos \theta_c, \sin \theta_c)$. If the matrices were not symmetric we would need to worry about where to multiply \mathbf{c}_j . Since our matrices have the j index second, we multiply on the right. Thus,

$$(\delta_{ij} - Q_{ij})c_j = \begin{pmatrix} 2\sin^2 \theta_c & -2\sin \theta_c \cos \theta_c \\ -2\sin \theta_c \cos \theta_c & 2\cos^2 \theta_c \end{pmatrix} \begin{pmatrix} \cos \theta_c \\ \sin \theta_c \end{pmatrix} = 2 \begin{pmatrix} \sin^2 \theta_c \cos \theta_c - \sin \theta_c \cos \theta_c \sin \theta_c \\ -\sin \theta_c \cos^2 \theta_c + \cos^2 \theta_c \sin \theta_c \end{pmatrix} = 0.$$

(d) Go to the book Web site and print out a full-sized version of figure 11.22. Cut out the hexagon, and fold along the edges. Where does the boundary go?

Most people figured out how to fold this properly (as far as I can tell what's proper). You can fold the figure on top of itself to produce an identical figure scaled by a factor of $1/2$ (so that the area is scaled by $1/4$). The boundary is hidden behind the still-existing boundaries and along the radial 'spokes' from the center to the vertices.

11.8 Minimizing Sequences.

(a) Find a minimizing sequence for the somewhat silly function g in figure 11.19.

We have $g(x) = x^2$ for $x \neq 0$ and $g(0) = 1$. A minimizing sequence for this function is a set $\{x_n\}$ such that $g(x_{n+1}) < g(x_n)$ and $\lim g(x_n) = g_0 = 0$. There are many such sequences. Any monotonic sequence converging to zero will do. One (rather slow) example is $x_n = 1/n$.

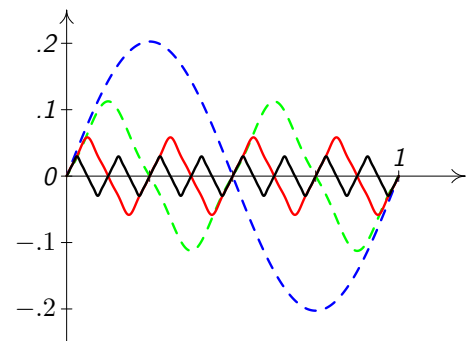
(b) Prove Microstructure Theorem 1. In particular,

- Show that zero is a lower bound for the energy \mathcal{F} .
- Construct a minimizing sequence of functions $y_n(x)$ for which $\lim \mathcal{F}[y_n] = 0$.
- Show that the second term of $\mathcal{F}[y]$ is zero only for $y(x) = 0$, which does not minimize \mathcal{F} .

We have $\mathcal{F}[y] = \int_0^1 (y'^2 - 1)^2 + y^2 dx = \int_0^1 F[y] dx$, with $y(0) = y(1) = 0$. We see first that zero is a lower bound for \mathcal{F} because F is the sum of squares and is therefore nonnegative. Thus, for all y , $\mathcal{F}[y] \geq 0$. Now we construct a sequence $\{y_n(x)\}$ which minimizes the energy. Let $y_n(x)$ be a triangle wave with period $1/2^n$ and amplitude $1/2^{n+1}$. This has slope ± 1 almost everywhere and $\int y_n^2 dx = 1/(3 \cdot 2^{n+1}) \rightarrow 0$. Some may object that this function is not differentiable at a lot of points. We could have defined

$$y_n(x) = \frac{1}{2^n} \frac{2}{\pi^2} \sum_{j=0}^n \frac{(-1)^j \sin(2^n [2j+1] 2\pi x)}{(2j+1)^2},$$

which is the Fourier sine series for this triangle wave. As $n \rightarrow \infty$ we add more terms and get closer and closer, but the function is always differentiable. Finally we show that $\mathcal{F}[y] > 0$ for all y . In particular, the second term is only zero if $y(x) = 0$ almost everywhere. If we take y to be continuous then $y(x) = 0$ for this to happen. But then $y' = 0$ so the first term gives $\int (y'^2 - 1)^2 = 1$. Thus there is no function $y(x)$ which minimizes \mathcal{F} , but we have found a minimizing sequence limiting to zero.



(c) Argue that the Young measure which describes minimizing sequences for the free energy in equation 11.28 is $\nu_x(S) = \frac{\delta}{2}(S-1) + \frac{\delta}{2}(S+1)$.

We have the Young measure $\nu_x(S)$ is the probability distribution for the slope S at points $x + \epsilon$ near x . For our limiting sequence, as $\mathcal{F} \rightarrow 0$ we need the slope to be ± 1 with increasing likelihood. Therefore, we have $\nu_x(S) = a(x)\delta(S - 1) + (1 - a(x))\delta(S + 1)$. We can now write $\langle y(x) \rangle$ as an integral, given the boundary condition $y(0) = 0$.

$$\langle y(x) \rangle = \left\langle \int_0^x dx' y'(x) \right\rangle = \int_0^x dx' \langle y'(x) \rangle = \int_0^x dx' (a(x)(+1) + (1 - a(x))(-1)) = \int_0^x dx' (2a(x) - 1).$$

To minimize \mathcal{F} we need also $\langle y(x) \rangle \rightarrow 0$ so that $2a(x) = 1$ everywhere. Thus, $\nu_x(s) = \frac{1}{2}\delta(S - 1) + \frac{1}{2}\delta(S + 1)$.

11.9 Snowflakes and Linear Stability.

(a) Generalize to 2D droplets. Find C ; find A .

The equations are simply modified to go from 3D to 2D:

$$\frac{dR}{dt} = -\frac{D\sigma}{\rho R^2}$$

$$R_c = \frac{\sigma T_v}{\rho L \Delta T}$$

Using 11.29, it is clear that:

$$C = \frac{D\sigma}{\rho}$$

and

$$A = \frac{C}{R_c^2} = \frac{D\rho}{\sigma} \left(L \frac{\Delta T}{T_v} \right)^2$$

(b) Find direction of motion for various curvatures

Positive curvature should go up and negative curvature should go down. A term of the form $-C\kappa^2$ would not capture this.

(c) Write $\hat{n} \cdot \frac{\partial x}{\partial t}$.

We have:

$$\hat{n} \cdot \frac{\partial x}{\partial t} = \frac{R}{\sqrt{R^2 + R_\theta^2}} \frac{\partial R}{\partial t}$$

(d) Find r_0 at which fingers start to grow.

Linearizing equation 11.29, we find:

$$\frac{\partial r_m}{\partial t} = r_m \left(B - \frac{2C}{r_0} \right) \left(\frac{m+2}{r_0^2} \right)$$

so, fingers start to grow when

$$r_0 = \frac{2C}{B}$$

12.2 Scaling and corrections to scaling.

[This solution is to an older version of the exercise. It may not correspond to the current version of the text.]

(a) Assuming ...

The exponent γ is defined such that for $T \rightarrow T_c$, $\chi(T) \sim |T - T_c|^{-\gamma}$. More specifically:

$$\gamma = - \lim_{T \rightarrow T_c} \frac{\log |\chi(T - T_c)|}{\log |T - T_c|} = 1.25.$$

(b) What is ...

The pair correlation function for a magnetic system goes like:

$$C(\vec{r}, T) \sim r^{-\tau} e^{-r/\xi},$$

where ξ is the correlation length $\xi(T) \sim |T - T_c|^{-\nu}$. Thus:

$$\tau = 1.026 \quad \text{and} \quad \nu = 0.59.$$

What about η ? At T_c , the correlation function becomes:

$$C(\vec{r}, T_c) \sim \frac{1}{r^{d-2+\eta}},$$

where d is the dimension of the space. Here $d = 3$, so $\eta = \tau - 1 = 0.026$.

12.3 Scaling and Coarsening.

[This solution is to an older version of the exercise. It may not correspond to the current version of the text.]

During coarsening ...

Recall that a system with a non-conserved order parameter has a length scale that goes like $L \sim t^{1/2}$. If $C(\vec{r}, T, t_0) = c(\vec{r})$, then at time $2t_0$ the length scale will have grown by a factor of $\sqrt{2}$. Since by assumption the correlation function is the same except for a rescaling of length:

$$C(\vec{r}, T, 2t_0) = c(\vec{r}/\sqrt{2}).$$

In writing the scaling form we note that the exponent $\omega = 0$, because $C(0, T, t)$ is time-independent (assuming it is finite and non-zero). And since:

$$C(\vec{r}, T, 2t_0) = \mathcal{C}(\vec{r}/(2t_0)^\rho, T) = \mathcal{C}(\vec{r}/(\sqrt{2}t_0)^\rho, T),$$

it must be that $\rho = 1/2$.

12.4 Bifurcation Theory and Phase Transitions.

(a) **Pitchfork.** Consider the differential equation in one variable $x(t)$ with one parameter μ

$$\dot{x} = \mu x - x^3.$$

Show that there is a bifurcation at $\mu_c = 0$, by showing that an initial condition with large $x(0)$ will evolve qualitatively differently at late times for $\mu > 0$ versus for $\mu < 0$.

A fixed point x^* has $\dot{x} = 0$. Therefore any fixed point of this map satisfies $\mu x^* = x^{*3}$. If $x^* \neq 0$ then $x^{*2} = \mu$. Thus, for $\mu < 0$ the only fixed point is at $x^*(\mu < 0) = 0$. If $\mu > 0$ then $x^*(\mu) = \pm\sqrt{\mu}$ are the stable fixed points. We can see this more clearly by looking at a picture. On the left is $\mu > 0$ and on the right is $\mu < 0$.



(b) Find the critical exponent β giving the size of the nonzero fixed points as a function of μ , $x^*(\mu) \propto (\mu - \mu_c)^\beta$.

We found above that $x^* = \sqrt{\mu}$. With $\mu_c = 0$ it is clear that $\beta = 1/2$.

(c) At what value λ_c does the differential equation

$$\dot{m} = \tanh(\lambda m) - m$$

have a bifurcation? Does the fixed point value $m^*(\lambda)$ behave as a power law $m^* \sim |\lambda - \lambda_c|^\beta$ near λ_c ? Does the value of β agree with that of the pitchfork bifurcation?

We can draw similar pictures as in part (a) to visualize this. The pictures basically show that we have stable fixed points for $d\dot{m}/dm < 0$. For small λ the slope of the \tanh is small and therefore $d\dot{m}/dm < 0$ everywhere. We have a single fixed point at the root $m = 0$. As we increase λ we will come to a point where

$$\frac{d\dot{m}}{dm} = 1 - \lambda \operatorname{sech}^2(\lambda m)$$

crosses zero at $m = 0$. Specifically, at $m = 0$, $\operatorname{sech}^2(\lambda m) = 1$ so $\lambda_c = 1$. In the case $\lambda > \lambda_c$ we now have an unstable fixed point at $m = 0$ and a pair of stable fixed points in either side. This is therefore qualitatively similar. Now let us determine the stable fixed points to leading order in $|\lambda - \lambda_c|$. We Taylor expand

$$\tanh(\lambda m) = \lambda m - \frac{(\lambda m)^3}{3} + \mathcal{O}(\lambda^5).$$

For $\dot{m} = 0$ we need $m = \tanh(\lambda m)$ so that $m(\lambda - 1) = m(\lambda - \lambda_c) = \frac{\lambda^3}{3} m^3$. Just as above this has a root at $m = 0$ but we are interested in the other roots,

$$m = \left(\frac{3}{\lambda^3} (\lambda - \lambda_c) \right)^{1/2} \approx \left(\frac{3}{\lambda_c^3} (\lambda - \lambda_c) \right)^{1/2}.$$

We see therefore that $m \propto (\lambda - \lambda_c)^\beta$ with $\beta = 1/2$.

12.5 Mean-field theory.

(a) At temperature $k_B T$, what is the value for $\langle s_j \rangle$ in equation 12.25, given m ? At what temperature T_c is the phase transition, in mean field theory? Argue from part (c) that $m \propto (T_c - T)^\beta$. Is this value for the critical exponent β correct for the Ising model in either two dimensions ($\beta = 1/8$) or three dimensions ($\beta \approx 1/3$)?

We have a result similar to earlier problems that $\langle s_j \rangle = \frac{1}{4J} \frac{1}{\beta} \frac{\partial}{\partial m} \log Z$ where the one-spin partition function is $Z = \sum_{s_j = \pm 1} \exp(4\beta J m s_j) = 2 \cosh(4m\beta J)$. We differentiate to find $\langle s_j \rangle = \tanh(4m\beta J)$. To get $m = \langle s_j \rangle$ we need $m = \tanh(4m\beta J)$, which is the same problem we solved in part (c) with $\lambda = 4\beta J$. Thus, $\lambda_c = 1$ and therefore $\beta_c = 1/4J$, or $T_c = 4J/k_B$. The scaling exponent is $\beta = 1/2$ just as in the pitchfork. Thus, it is grossly wrong for $d = 2$, and a little better for $d = 3$. When we get to $d = 4$ and higher, we can expect much better agreement.

(b) [No answer available yet].

(c) [No answer available yet].

12.7 Renormalization Group Trajectories.

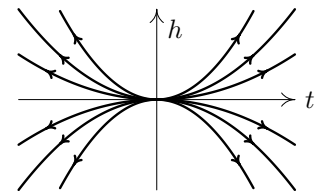
(a) Which diagram from figure 12.23 has curves consistent with this flow, for $b > a > 0$?

D. We begin by looking at small ϵ . We have $\Delta h = b\epsilon h$ and $\Delta t = a\epsilon t$. If we let $\epsilon = \Delta n$ where n is the coarse-graining coordinate which increases along a trajectory, then $\Delta h/\epsilon \rightarrow dh/dn$. We can now solve the differential equations to see

$$h(n) = h_0 e^{bn} \quad t(n) = t_0 e^{an}.$$

We can solve both of these for n and equate to find

$$h = h_0 \left(\frac{t}{t_0} \right)^{\frac{b}{a}}.$$



If $b > a > 0$ then this will be concave away from the t -axis, so **D** is the correct choice. This is what we might have guessed, since $b > a$ means that h is a more relevant eigendirection so the trajectories should curve more in that direction.

(b) Suppose $M(t, h)$ is known at $t = t_1$, the line of filled circles in the various figures in part (a). Give a formula for $M(2t_1, h)$ (open circles) in terms of $M(t_1, h')$.

We have M under coarse-graining scales as $(1 + c\epsilon)M(t, h) = M((1 + a\epsilon)t, (1 + b\epsilon)h)$. Using the same method as above we can convert this to a differential equation $dM/dn = cM$. Again we solve

$$M(t(n), h(n)) = M(t(0), h(0))e^{cn}.$$

Now take $t(0) = t_1$ and $h(0) = h'$. Then our results from part (a) give $t(n) = 2t_1$ so that $e^{an} = 2$, or $n = \frac{1}{a} \log 2$. We let $h(n) = h$ and now we find

$$M(2t_1, h) = M(t_1, h')e^{cn} = 2^{\frac{c}{a}} M(t_1, h').$$

In slightly more detail, we can label at any point in the (t, h) space by the trajectory it's on and a coordinate along that flow. Let us set $t(n = 0) = t_0 = 1$. We can therefore uniquely identify any trajectory (at least the half with $t > 0$) by

giving $h(n=0) = h_0$. For (t, h) we therefore have $t = t_0 e^{an}$ so that $n = \frac{1}{a} \log(t/t_0) = \frac{1}{a} \log t$. Then the trajectory has $h_0 = h/t^{b/a}$. Using our initial result from above, we can define \mathcal{M} as follows,

$$\mathcal{M}(h) \equiv M(t=1, h) \Rightarrow M(t, h) = t^{\frac{c}{a}} \mathcal{M}\left(\frac{h}{t^{\frac{b}{a}}}\right).$$

12.8 Superconductivity and the Renormalization Group.

(a) Taylor expand T_c about $V = 0^+$. Guess or show the value of all the terms in the Taylor series. Can we expect to explain superconductivity at positive temperatures by perturbing in powers of V ?

We Taylor expand $T_c(V) = \alpha e^{-1/Vg(\epsilon)}$ where $\alpha = 1.764\hbar\omega_D$. We differentiate

$$\frac{\partial T_c}{\partial V} = \alpha \left(e^{-1/Vg} \right) \left(\frac{1}{gV^2} \right)$$

$$\frac{\partial^2 T_c}{\partial V^2} = \alpha \left(e^{-1/Vg} \right) \left(\frac{1}{g^2 V^4} - \frac{2}{gV^3} \right).$$

We can therefore write $T_c(V) \stackrel{?}{\approx} \alpha e^{-1/Vg} P(1/V)$ where P is a polynomial. Although the argument of the polynomial diverges as $V \rightarrow 0$, the exponential converges to zero much faster, so this expansion vanishes at all orders. We won't be explaining superconductivity with a power series perturbation in V .

(b) True or false?

(T) (F) For $V > 0$ (attractive), the interactions get stronger with coarse-graining.

True. According to 12.35, $dV/dn > 0$ always, since $b > 0$. Therefore the interaction strength V gets farther from zero when coarse graining when V is positive.

(T) (F) For $V < 0$ (repulsive), coarse graining leads us back to the free Fermi gas, explaining why the Fermi gas describes metals.

True. We see that the interaction strength gets weaker (closer to zero) under coarse graining since $dV/dn > 0$ still. Visually we can see the flows leading to the high- T fermi gas.

(T) (F) Temperature is an irrelevant variable, but dangerous.

False. The critical point is unstable in the direction of temperature. If we start on the critical line and change the temperature slightly in one direction or the other, we have vastly different results (and a phase transition). Irrelevant variables must be attracted to a fixed point (i.e. stable).

(T) (F) The scaling variable $x = TV^{-\beta\delta}$ is unchanged by the coarse-graining, where β and δ are universal critical exponents: hence x labels the progress along the curves in figure 12.25.

False. *This statement starts out promising, since x is indeed unchanged by coarse graining ($dx/dn = 0$). But this means precisely the opposite: x can't be used to measure progress along a curve, since it is constant along the curve.*

(T) (F) The scaling variable $y = Te^{a/bV}$ is unchanged by the coarse-graining, so each curve in figure 12.25 has a fixed value for y .

True. *We differentiate*

$$\frac{dy}{dn} = \left(\frac{dT}{dn}\right) e^{\frac{a}{bV}} + T \frac{d}{dn} \left(e^{\frac{a}{bV}}\right) = \frac{dT}{dn} e^{\frac{a}{bV}} + T \frac{a}{b} \left(-\frac{1}{V^2}\right) \frac{dV}{dn} e^{\frac{a}{bV}}.$$

We recognize $dT/dn = aT$ and $dV/dn = bV^2$ so that this sum vanishes. Thus, y is also unchanged under coarse graining and therefore each curve has a fixed value of y .

(c) What value for a/b must they calculate in order to get the BCS transition temperature from this renormalization group? What is the nature of the scaling variable along $T_c(V)$?

Since $T_c = \alpha e^{-1/Vg}$, we can write the constant $\alpha = 1.764\hbar\omega_D = T_c e^{-1/Vg}$. We can now identify this with $y = Te^{a/bV}$ to find

$$\frac{a}{b} = \frac{1}{g} \quad y = 1.764\hbar\omega_D.$$

12.10 Renormalization Group and the Central Limit Theorem (Short).

See solution to exercise 12.11.

12.11 Renormalization Group and the Central Limit Theorem (Long).

(a) Argue that if $\rho(x)$ is the probability that a random variable has value x , that the probability distribution of the sum of two random variables drawn from this distribution is the convolution

$$C[\rho](x) = (\rho * \rho)(x) = \int_{-\infty}^{\infty} \rho(x-y)\rho(y)dy.$$

We are given a probability distribution ρ and two random variables from it, y and y' . We want the probability distribution of the sum $x = y + y'$, $C[\rho]$. The probabilities must therefore satisfy

$$C[\rho](x) dx = \int_{y+y'=x} \rho(y)\rho(y') dy dy'.$$

For any x and y we can write $y' = x - y$ to satisfy the condition, and therefore $dy' = dx$ for a y constant. This gives

$$C[\rho](x) dx = \left(\int_{-\infty}^{\infty} \rho(y)\rho(x-y) dy \right) dx.$$

Dividing out the dx , we see that this is the definition of a convolution, $C[\rho](x) = (\rho * \rho)(x)$.

(b) Show that ρ is normalized, so is $S_{\sqrt{2}}[\rho]$. Show that the Fourier transform

$$\mathcal{F}[S_{\sqrt{2}}[\rho]](k) = \tilde{\rho}(k/\sqrt{2}).$$

Now we consider the rescaling operation (more generally), $S_a[\rho](x) = a\rho(ax)$. The Fourier transform of this is

$$\mathcal{F}[S_a[\rho]](k) = \int_{-\infty}^{\infty} dx e^{ikx} S_a[\rho](x) = \int_{-\infty}^{\infty} dx e^{ikx} a\rho(ax) = \int_{-\infty}^{\infty} d[ax] e^{i\frac{k}{a}(ax)} \rho(ax) = \mathcal{F}[\rho]\left(\frac{k}{a}\right) = \tilde{\rho}\left(\frac{k}{a}\right).$$

(c) Show that the Gaussian distribution

$$\rho^*(x) = (1/\sqrt{2\pi}\sigma) \exp(-x^2/2\sigma^2)$$

is indeed a fixed point in function space under the operation T . You can do this either by direct integration, or by using the known properties of the Gaussian under convolution.

With $T[\rho] = S_{\sqrt{2}}[C[\rho]]$, we have $T[\rho^*](x) = \sqrt{2} \int_{-\infty}^{\infty} \rho^*(\sqrt{2}x - y) \rho^*(y) dy$. We immediately substitute $y = \sqrt{2}u$. Now

$$\begin{aligned} T[\rho^*](x) &= 2 \int_{-\infty}^{\infty} \rho^*(\sqrt{2}(x - u)) \rho^*(\sqrt{2}u) du = \frac{1}{\pi\sigma^2} \int_{-\infty}^{\infty} e^{-(x^2+u^2)/\sigma^2} e^{-u^2/\sigma^2} du \\ &= \frac{1}{\pi\sigma^2} e^{-x^2/2\sigma^2} \int_{-\infty}^{\infty} e^{-(u-\frac{x}{2})^2/2(\sigma/2)^2} du = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} = \rho^*(x). \end{aligned}$$

We see that Gaussians are a fixed point of T .

(d) Use equations 12.41 and 12.43 to show that

$$\mathcal{F}[T[\rho]](k) = \tilde{T}[\tilde{\rho}](k) = \tilde{\rho}(k/\sqrt{2})^2.$$

Calculate the Fourier transform of the fixed point $\tilde{\rho}^*(k)$. Using equation 12.46, show that $\tilde{\rho}^*(k)$ is a fixed point in Fourier space under our coarsening operator \tilde{T} .

We start by expanding $\mathcal{F}[T[\rho]](k) = \mathcal{F}[S_{\sqrt{2}}[C[\rho]]](k)$. Taking $C[\rho]$ for ρ in (12.43) we can write $\mathcal{F}[S_{\sqrt{2}}[C[\rho]]](k) = \mathcal{F}[C[\rho]](k/\sqrt{2})$. Now (12.41) gives $\mathcal{F}[C[\rho]](k/\sqrt{2}) = (\tilde{\rho}(k/\sqrt{2}))^2$. Thus,

$$\mathcal{F}[T[\rho]](k) = \tilde{T}[\tilde{\rho}](k) = \tilde{\rho}(k/\sqrt{2})^2.$$

Now we Fourier transform our Gaussian,

$$\tilde{\rho}^*(k) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} dx e^{ikx} e^{-x^2/2\sigma^2} = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} dx e^{\frac{1}{2}\sigma^2 k^2} e^{-(x^2 - 2\sigma^2 ikx - \sigma^4 k^2)/2\sigma^2} = e^{-\frac{1}{2}\sigma^2 k^2},$$

where we've gone through the usual drill of completing the square and magically hand-waved over some messy complex analysis which doesn't end up mattering in the end. We see that Gaussians are indeed eigenfunctions of the Fourier transform. Now we consider

$$\tilde{T}[\tilde{\rho}^*] = \left[\tilde{\rho}^* \left(\frac{k}{\sqrt{2}} \right) \right]^2 = e^{-2\frac{1}{2}\sigma^2(k/\sqrt{2})^2} = e^{-\frac{1}{2}\sigma^2 k^2} = \tilde{\rho}^*[k].$$

(e) Show using equations 12.46 and 12.47 that the transformations of the eigenfunctions satisfy

$$\tilde{f}_n(k) = (2/\lambda_n)\tilde{\rho}^*(k/\sqrt{2})\tilde{f}_n(k/\sqrt{2}).$$

We seek eigenfunctions f_n and eigenvalues λ_n of the derivative of the mapping T . That is, $\tilde{T}[\tilde{\rho}^* + \epsilon \tilde{f}_n] = \tilde{\rho}^* + \lambda_n \epsilon \tilde{f}_n + \mathcal{O}(\epsilon^2)$. Using our previous results we can expand \tilde{T} to find

$$\tilde{T}[\tilde{\rho}^* + \epsilon \tilde{f}_n](k) = \left[(\tilde{\rho}^*(k/\sqrt{2}) + \epsilon \tilde{f}_n(k/\sqrt{2})) \right]^2 = \tilde{\rho}^*(k) + 2\epsilon \tilde{\rho}^*(k/\sqrt{2})\tilde{f}_n(k/\sqrt{2}) + \mathcal{O}(\epsilon^2).$$

Rearranging terms, we need

$$\tilde{f}_n(k) = \frac{2}{\lambda_n} \tilde{\rho}^* \left(\frac{k}{\sqrt{2}} \right) \tilde{f}_n \left(\frac{k}{\sqrt{2}} \right).$$

(f) Show that $\tilde{f}_n(k) = (ik)^n \tilde{\rho}^*(k)$ is the Fourier transform of an eigenfunction. What is the eigenvalue λ_n ?

If f_n is to be an eigenfunction then we need $\tilde{f}_n(k) = (2/\lambda_n)\tilde{\rho}^*(k/\sqrt{2})\tilde{f}_n(k/\sqrt{2})$. Substituting $\tilde{f}_n(k) = (ik)^n \tilde{\rho}^*(k)$ and recognizing that $(\tilde{\rho}^*(k/\sqrt{2}))^2 = \tilde{\rho}^*(k)$, we have

$$\tilde{f}_n(k) = (ik)^n \tilde{\rho}^*(k) = \left(\frac{2}{\lambda_n} \right) \tilde{\rho}^* \left(\frac{k}{\sqrt{2}} \right) \tilde{\rho}^* \left(\frac{k}{\sqrt{2}} \right) \left(\frac{ik}{\sqrt{2}} \right)^n.$$

Cancelling the $\tilde{\rho}^*(k)$ and $(ik)^n$ gives

$$\lambda_n = \frac{2}{\sqrt{2}^n} = 2^{\frac{1}{2}(2-n)}.$$

(g) The eigenfunction $f_0(x)$ with the biggest eigenvalue corresponds to an unphysical perturbation: why? The next two eigenfunctions f_1 and f_2 have important physical interpretations. Show that $\rho^* + \epsilon f_1$ to lowest order is equivalent to a shift in the mean of ρ and $\rho^* + \epsilon f_2$ is a shift in the standard deviation σ of ρ^* .

We note that $f_0 = \rho^*$. But since f_n is merely a perturbation which we add to ρ^* , we need $\int \rho(x)dx = \int (\rho^*(x) + f_n(x))dx = 1 + \int_{-\infty}^{\infty} f_n(x)dx = 1$ so that $\int_{-\infty}^{\infty} f_n(x)dx = 0$. Thus, f_0 is unphysical since it is normalized to one, but we need our perturbations to be normalized to zero.

More generally we can use the property of Fourier transforms $\mathcal{F}[\frac{\partial}{\partial x} f(x)] = -ik\tilde{f}(k)$ to see that $f_n(x) = (-\frac{\partial}{\partial x}) \rho^*(x) = -(-1/\sigma)^n \text{He}_n(x/\sigma) \rho^*(x)$, where $\text{He}_n(x)$ is the n^{th} Hermite polynomial, as found on Mathworld (where it's considered a 'rarely used alternate definition'). Thus, $f_1(x) = \frac{x}{\sigma^2} \rho^*(x)$. Consider a shift in the mean of ρ^* . That is,

$$\rho'(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\epsilon)^2/2\sigma^2} \approx \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} e^{\epsilon x/\sigma^2} \approx \rho^*(x) \left(1 + \epsilon \frac{x}{\sigma^2}\right) = \rho^*(x) + \epsilon f_1(x).$$

So the perturbation f_1 with eigenvalue $\lambda_1 = \sqrt{2}$ corresponds to a shift in the mean. Now consider a shift in the standard deviation, $\sigma \rightarrow \sigma + \epsilon/\sigma$. If you think this is an odd choice, we can look at the variance, $\sigma^2 \rightarrow \sigma^2 + 2\epsilon$ to first order, and see that it makes a little more sense now. In this case, it is much easier to work in Fourier space again. Thus,

$$\tilde{\rho}''(k) \approx e^{-\frac{1}{2}(\sigma^2+2\epsilon)k^2} = e^{-\frac{1}{2}\sigma^2 k^2} e^{-\epsilon k^2} \approx \tilde{\rho}''(k) (1 - \epsilon k^2) = \tilde{\rho}''(k) + \tilde{f}_2(k).$$

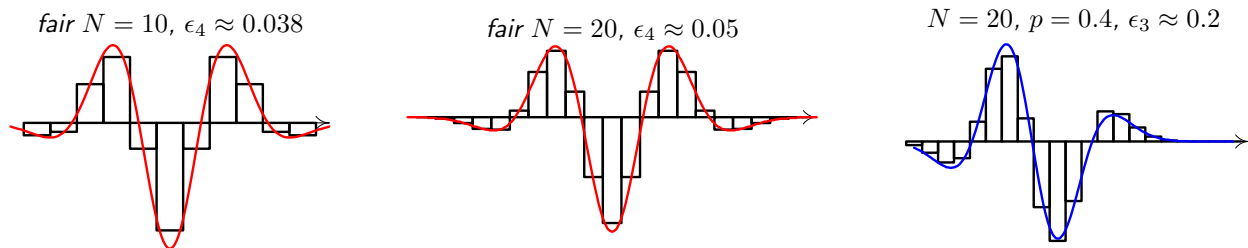
So the second perturbation ($\lambda_2 = 1$) is a shift in the variance. This is important, since we need any eigenvalues ≥ 1 to be trivial.

(h) Plot the difference between the binomial distribution of N coin flips and a Gaussian of the same mean and width, for $N = 10$ and $N = 20$. Does it approach one of the eigenfunctions f_3 or f_4 ?

(i) Why didn't a perturbation along $f_3(x)$ dominate the asymptotics? What symmetry forced $\epsilon_3 = 0$? Should flips of a biased coin break this symmetry?

We plot the difference between the binomial distribution and the Gaussian distribution for $N = 10$ and 20 in the first two plots below. We see that these residues resemble f_4 very well.

What happened to f_3 ? The binomial and Gaussian distributions are both symmetric about their mean and so the difference must also be. But f_3 is an odd function and thus is antisymmetric. If we throw an unfair coin, the mean and standard deviation will both shift. If we shift our Gaussian accordingly, we should expect that an f_3 term will remain, as is seen in the right-most graph, which has a probability of $2/5$.



We could have made prettier graphs by using a "continuumized" version of the binomial distribution constructed from Γ functions.

12.13 Hysteresis Model: Scaling and Exponent Equalities.

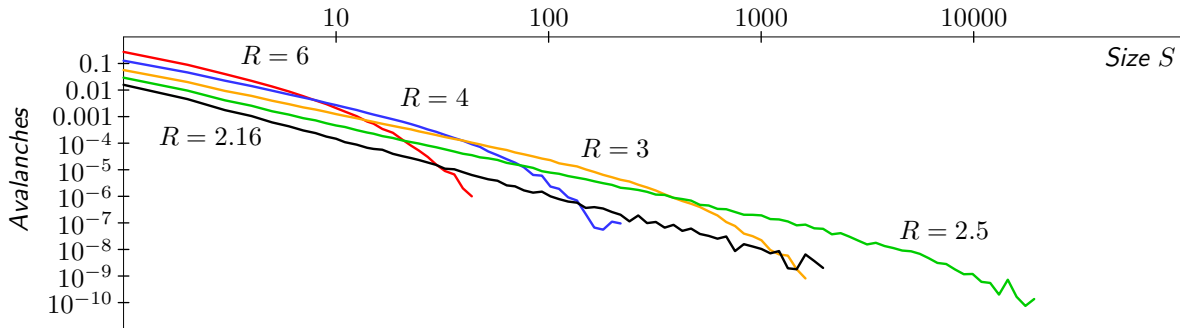
(a) What fraction of the system is flipped by the one largest avalanche, in your simulation? Compare this with the hysteresis curve at $R = 2.4 > R_c$. Does it have a similar bug jump, or is it continuous?

At $R = R_c \approx 2.16$ you should find a large jump that flips anywhere from 70% to 90% of the spins at once. On the other hand, for $R = 2.4 > R_c$ you should see several to many smaller avalanches making a more continuous transition.

(b) From your plot, measure this exponent combination from your simulation. It should be close to two. Is your estimate larger or smaller than two?

Most people found a power law $\tau + \sigma\beta\delta \approx 2.05$.

(c) Copy and edit your avalanche size distribution files, removing the data after the first bin with zero avalanches in it. Start up a graphics program and plot the curves on a log-log plot: they should look like power laws for small S and cut off at larger S .



(d) What are A and B in this equation for the scaling form given by equation 12.65?

To find A and B , we write the coarse-graining equation, $D(S, R - R_c) = AD(BS, 2(R - R_c))$. We can now substitute 12.65 into both sides, so that

$$S^{-(\tau+\sigma\beta\delta)} \mathcal{D}(S(R - R_c)^{1/\sigma}) = A(BS)^{-\tau(\sigma\beta\delta)} \mathcal{D}(BS2^{1/\sigma}(R - R_c)^{1/\sigma}).$$

For equality at any R we therefore need $AB^{-(\tau+\sigma\beta\delta)} = 1$ and $2^{1/\sigma}B = 1$. Thus,

$$B = 2^{-1/\sigma} \quad A = 2^{-(\tau/\sigma+\beta\delta)}.$$

(e) Multiply the vertical axis of each curve by $S^{\tau+\sigma\beta\delta}$. This then should give four curves $\mathcal{D}(S(R - R_c)^{1/\sigma})$ which are (on a log-log plot) roughly the same shape, just shifted sideways horizontally. Measure the peak of each curve. Make a table with columns R , S_{peak} , and $R - R_c$. Do a log-log plot of $R - R_c$ versus S_{peak} and estimate σ in the expected power law $S_{\text{peak}} \sim (R - R_c)^{-1/\sigma}$.

The plots should be rather similar, only shifted. The $R = 4$ and $R = 6$ plots are the most outlying. Most people found an exponent σ between $1/3$ and $4/5$.

(f) Do a scaling collapse: plot $S^{\tau+\sigma\beta\delta} D(S, R)$ versus $(R - R_c)^{1/\sigma} S$ for the avalanche size distributions with $R > R_c$. How well do they collapse onto a single curve?

Using $\sigma = 4/5$ and $\tau + \sigma\beta\delta = 2$, we find scaling collapse.

