PHYSICS 210A : STATISTICAL PHYSICS
HW ASSIGNMENT #1 SOLUTIONS

(1) Prove that for $x \geq 0$ and $y \geq 0$ that

$$(x - y)\,(\ln x - \ln y) \geq 0 \ .$$

Solution: Trivial. Both $f(x) = x$ and $g(x) = \ln x$ are strictly increasing functions on the interval $(0, \infty)$. Hence $\ln x < \ln y$ if $0 < x < y$, and $\ln y < \ln x$ if $0 < y < x$. Thus, $(x - y)\big(\ln x - \ln y\big) \geq 0$.

(2) A *Markov chain* is a process which describes transitions of a discrete stochastic variable occurring at discrete times. Let $P_i(t)$ be the probability that the system is in state i at time t. The evolution equation is

$$P_i(t+1) = \sum_j Q_{ij}\, P_j(t) \ .$$

The *transition matrix* Q_{ij} satisfies $\sum_i Q_{ij} = 1$ so that the total probability $\sum_i P_i(t)$ is conserved. The element Q_{ij} is the *conditional probability* that for the system to evolve to state i given that it is in state j. Now consider a group of Physics graduate students consisting of three theorists and four experimentalists. Within each group, the students are to be regarded as indistinguishable. Together, the students rent two apartments, A and B. Initially the three theorists live in A and the four experimentalists live in B. Each month, a random occupant of A and a random occupant of B exchange domiciles. Compute the transition matrix Q_{ij} for this Markov chain, and compute the average fraction of the time that B contains two theorists and two experimentalists, averaged over the effectively infinite time it takes the students to get their degrees. *Hint:* Q is a 4×4 matrix.

Solution: There are four states available: Now let's compute the transition probabilities.

$\lvert\, j \,\rangle$	room A	room B	g_j^{A}	g_j^{B}	g_j^{TOT}
$\lvert\, 1 \,\rangle$	TTT	EEEE	1	1	1
$\lvert\, 2 \,\rangle$	TTE	EEET	3	4	12
$\lvert\, 3 \,\rangle$	TEE	EETT	3	6	18
$\lvert\, 4 \,\rangle$	EEE	ETTT	1	4	4

Table 1: States and their degeneracies.

First, we compute the transition probabilities out of state $\lvert\, 1 \,\rangle$, *i.e.* the matrix elements Q_{j1}. Clearly $Q_{21} = 1$ since we must exchange a theorist (T) for an experimentalist (E). All the other probabilities are zero: $Q_{11} = Q_{31} = Q_{41} = 0$. For transitions out of state $\lvert\, 2 \,\rangle$, the nonzero elements are

$$Q_{12} = \tfrac{1}{4} \times \tfrac{1}{3} = \tfrac{1}{12} \quad , \quad Q_{22} = \tfrac{3}{4} \times \tfrac{1}{3} + \tfrac{1}{4} \times \tfrac{2}{3} = \tfrac{5}{12} \quad , \quad Q_{32} = \tfrac{1}{2} \ .$$

To compute Q_{12}, we must choose the experimentalist from room A (probability $\frac{1}{3}$) with the theorist from room B (probability $\frac{1}{4}$). For Q_{22}, we can either choose E from A and one of the E's from B, or one of the T's from A and the T from B. This explains the intermediate steps written above. For transitions out of state $|\,3\,\rangle$, the nonzero elements are then

$$Q_{23} = \tfrac{1}{3} \quad , \quad Q_{33} = \tfrac{1}{2} \quad , \quad Q_{43} = \tfrac{1}{6} \ .$$

Finally, for transitions out of state $|\,4\,\rangle$, the nonzero elements are

$$Q_{34} = \tfrac{3}{4} \quad , \quad Q_{44} = \tfrac{1}{4} \ .$$

The full transition matrix is then

$$Q = \begin{pmatrix} 0 & \frac{1}{12} & 0 & 0 \\ 1 & \frac{5}{12} & \frac{1}{3} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & \frac{3}{4} \\ 0 & 0 & \frac{1}{6} & \frac{1}{4} \end{pmatrix} \ .$$

Note that $\sum_i Q_{ij} = 1$ for all $j = 1, 2, 3, 4$. This guarantees that $\phi^{(1)} = (1\,,\,1\,,\,1\,,\,1)$ is a left eigenvector of Q with eigenvalue 1. The corresponding right eigenvector is obtained by setting $Q_{ij}\,\psi^{(1)}_j = \psi^{(1)}_i$. Simultaneously solving these four equations and normalizing so that $\sum_j \psi^{(1)}_j = 1$, we easily obtain

$$\psi^{(1)} = \frac{1}{35} \begin{pmatrix} 1 \\ 12 \\ 18 \\ 4 \end{pmatrix} \ .$$

This is the state we converge to after repeated application of the transition matrix Q. If we decompose $Q = \sum_{\alpha=1}^{4} \lambda_\alpha \,|\,\psi^{(\alpha)}\,\rangle\langle\,\phi^{(\alpha)}\,|$, then in the limit $t \to \infty$ we have $Q^t \approx |\,\psi^{(1)}\,\rangle\langle\,\phi^{(1)}\,|$, where $\lambda_1 = 1$, since the remaining eigenvalues are all less than 1 in magnitude[1]. Thus, Q^t acts as a *projector* onto the state $|\,\psi^{(1)}\,\rangle$. Whatever the initial set of probabilities $P_j(t=0)$, we must have $\langle\,\phi^{(1)}\,|\,P(0)\,\rangle = \sum_j P_j(0) = 1$. Therefore, $\lim_{t\to\infty} P_j(t) = \psi^{(1)}_j$, and we find $P_3(\infty) = \frac{18}{35}$. Note that the equilibrium distribution satisfies detailed balance:

$$\psi^{(1)}_j = \frac{g^{\rm TOT}_j}{\sum_l g^{\rm TOT}_l} \ .$$

(3) Consider a q-state generalization of the Kac ring model in which $\mathbb{Z}_q$ spins rotate around an N-site ring which contains a fraction $x = N_{\rm F}/N$ of flippers on its links. Each flipper cyclically rotates the spin values: $1 \to 2 \to 3 \to \cdots \to q \to 1$ (hence the clock model symmetry $\mathbb{Z}_q$).

[1]One can check that $\lambda_1 = 1$, $\lambda_2 = \frac{5}{12}$, $\lambda_3 = -\frac{1}{4}$. and $\lambda_4 = 0$.

(a) What is the Poincare recurrence time?

(b) Make the *Stosszahlansatz*, *i.e.* assume the spin flips are stochastic random processes. Then one has

$$P_\sigma(t+1) = (1-x)\,P_\sigma(t) + x\,P_{\sigma-1}(t) \ ,$$

where $P_0 \equiv P_q$. This defines a Markov chain

$$P_\sigma(t+1) = Q_{\sigma\sigma'}\,P_{\sigma'}(t) \ .$$

Decompose the transition matrix Q into its eigenvectors. *Hint:* The matrix may be diagonalized by a simple Fourier transform.

(c) The eigenvalues of Q may be written as $\lambda_\alpha = e^{-1/\tau_\alpha}\,e^{-i\delta_\alpha}$, where τ_α is a relaxation time and δ_α is a phase. Find the spectrum of relaxation times. What is the longest finite relaxation time?

(d) Suppose all the spins are initially in the state $\sigma = q$. Write down an expression for $P_\sigma(t)$ for all subsequent times $t \in \mathbb{Z}^+$. Plot your results for different values of x and q.

Hint: It may be helpful to study carefully the solution to problem 5.1 (*i.e.* problem 1 of assignment 5) from F08 Physics 140A. You can access this through the link to the 140B website on the 210A course web page.

Solution:

(a) The recurrence time is $\tau = qN/\mathsf{gcd}(N_{\rm F}, q)$, where $\mathsf{gcd}(N_{\rm F}, q)$ is the greatest common divisor of $N_{\rm F}$ and q. After τ steps, which is to say $q/\mathsf{gcd}(N_{\rm F}, q)$ cycles around the ring, each spin will have visited $qN_{\rm F}/\mathsf{gcd}(N_{\rm F}, q)$ flippers. This is necessarily an integer multiple of q, which means that each spin will have mate $N_{\rm F}/\mathsf{gcd}(N_{\rm F}, q)$ complete cycles of its internal $\mathbb{Z}_q$ clock.

(b) We have

$$Q_{\sigma\sigma'} = (1-x)\,\widetilde{\delta}_{\sigma,\sigma'} + x\,\widetilde{\delta}_{\sigma,\sigma'+1} \ ,$$

where

$$\widetilde{\delta}_{ij} = \begin{cases} 1 & \text{if } i = j \bmod q \\ 0 & \text{otherwise.} \end{cases}$$

Q is known as a *circulant matrix*, which is to say it satisfies $Q_{\sigma\sigma'} = Q(\sigma - \sigma' \bmod q)$. A circulant matrix of rank q has only q independent entries. Such a matrix may be brought to diagonal form by a unitary transformation: $Q = U\,\widehat{Q}\,U^\dagger$,[2] where $U_{\sigma k} = \frac{1}{\sqrt{q}}\,e^{2\pi i k\sigma/q}$ and $\widehat{Q}_{kk'} \equiv \widehat{Q}(k)\,\widetilde{\delta}_{kk'}$ with

$$\widehat{Q}(k) = \sum_{n=1}^{q} Q(\mu)\,e^{-2\pi i k\mu/q} \ . \qquad (1)$$

Since $Q(\mu) = (1-x)\,\widetilde{\delta}_{\mu,0} + x\,\widetilde{\delta}_{\mu,1}$, we have

$$\widehat{Q}(k) = 1 - x + x\,e^{-2\pi i k/q} \ .$$

[2] There was some discussion of the details on the web forum pages for Physics 210A this past week.

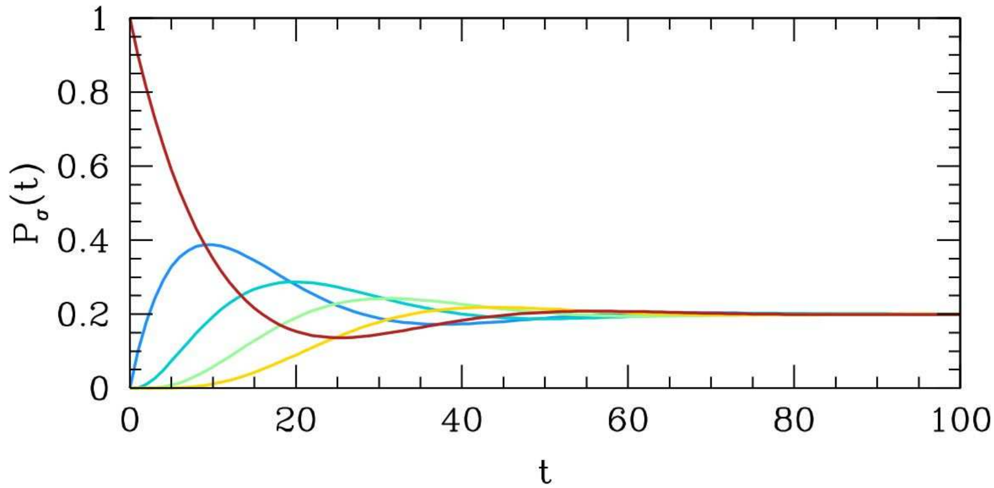

Figure 1: Behavior of $P_\sigma(t)$ for $q = 5$ and $x = 0.1$ within the *Stosszahlansatz* with initial conditions $P_\sigma(0) = \delta_{\sigma,q}$. Note that at large times the probabilities all converge to $\lim_{t\to\infty} P_\sigma(t) = q^{-1}$.

(c) In the polar representation, we have $\widehat{Q}(k) = e^{-1/\tau_k(x)}\, e^{-i\delta_k(x)}$, where

$$\tau_k(x) = -\frac{2}{\ln\Big[1 - 2x(1-x)\big(1-\cos(2\pi k/q)\big)\Big]}$$

and

$$\delta_k(x) = \tan^{-1}\left(\frac{x\,\sin(2\pi k/q)}{1 - x + x\,\cos(2\pi k/q)}\right)\ .$$

Note that $\tau_q = \infty$, because the total probability is conserved by the Markov process. The longest finite relaxation time is $\tau_1 = \tau_{q-1}$.

(d) Given the initial conditions $P_\sigma(0) = \delta_{\sigma,q}$, we have

$$\begin{aligned} P_\sigma(t) &= \big(Q^t\big)_{\sigma\sigma'}\, P_{\sigma'}(0) \\ &= \frac{1}{q}\sum_{k=1}^{q} U_{\sigma k}\, \widehat{Q}^t(k)\, U^*_{\sigma' k}\, P_{\sigma'}(0) \\ &= \frac{1}{q}\sum_{k=1}^{q} e^{-t/\tau_k}\, e^{-it\delta_k}\, e^{2\pi i\sigma k/q}\ . \end{aligned}$$

We can combine the terms in the k sum by pairing k with $q-k$, since $\tau_{q-k} = \tau_k$ and $\delta_{q-k} = -\delta_k$. We should however consider separately the cases $k = q$ and, if q is even, $k = \frac{1}{2}q$, since for those values of k we have $\widehat{Q}(k)$ is real.

If q is even, then $\widehat{Q}\big(k = \frac{1}{2}q\big) = 1 - 2x$. We then have

$$P_\sigma(t) = \frac{1}{q} + \frac{(-1)^\sigma}{q}\,(1-2x)^t + \frac{2}{q}\sum_{k=1}^{\frac{q}{2}-1} e^{-t/\tau_k(x)}\, \cos\!\left(\frac{2\pi\sigma k}{q} - t\,\delta_k(x)\right)\ .$$

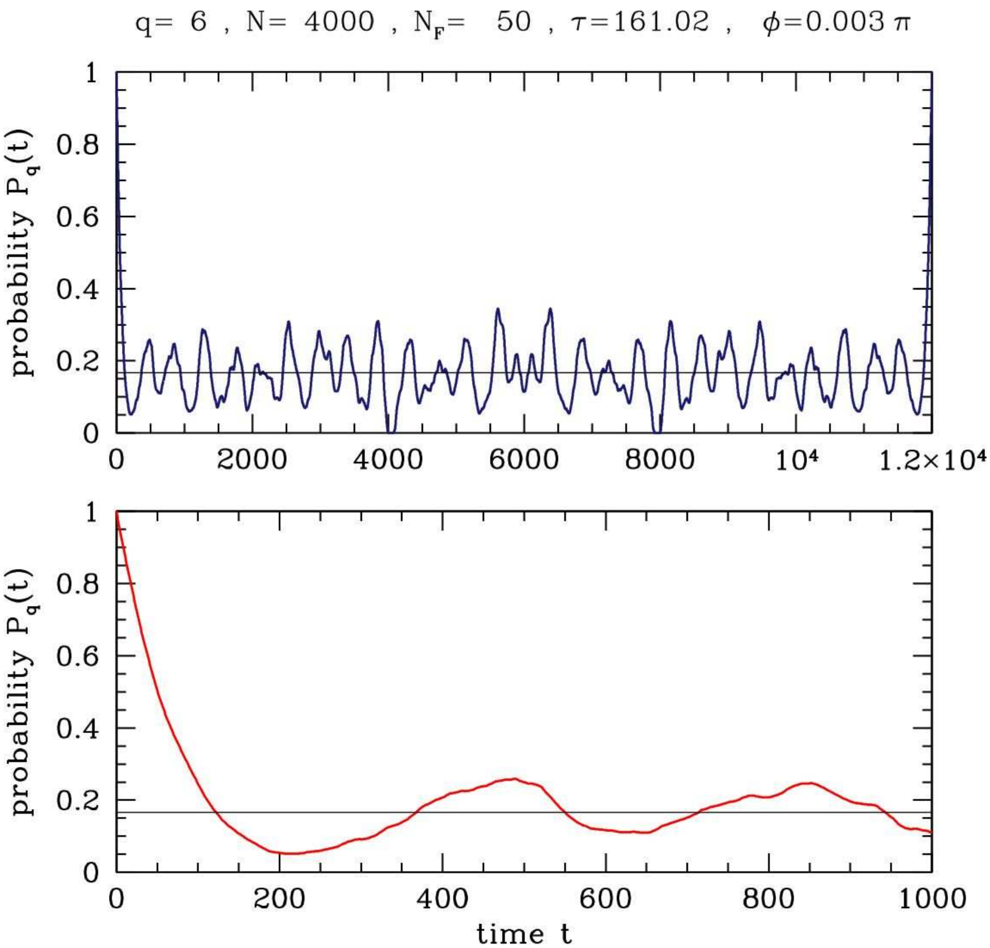

Figure 2: Evolution of the initial distribution $P_\sigma(0) = \delta_{\sigma,q}$ for the $\mathbb{Z}_q$ Kac ring model for $q = 6$, from a direct numerical simulation of the model.

If q is odd, then

$$P_\sigma(t) = \frac{1}{q} + \frac{2}{q} \sum_{k=1}^{\frac{q-1}{2}} e^{-t/\tau_k(x)} \cos\!\left(\frac{2\pi\sigma k}{q} - t\,\delta_k(x)\right) .$$

(e) See fig. 2.

(4) A ball of mass m executes perfect one-dimensional motion along the symmetry axis of a piston. Above the ball lies a mobile piston head of mass M which slides frictionlessly inside the piston. Both the ball and piston head execute ballistic motion, with two types of collision possible: (i) the ball may bounce off the floor, which is assumed to be infinitely massive and fixed in space, and (ii) the ball and piston head may engage in a one-dimensional elastic collision. The Hamiltonian is

$$H = \frac{P^2}{2M} + \frac{p^2}{2m} + MgX + mgx \ ,$$

where X is the height of the piston head and x the height of the ball. Another quantity is conserved by the dynamics: $\Theta(X - x)$. *I.e.*, the ball always is below the piston head.

(a) Choose an arbitrary length scale L, and then energy scale $E_0 = MgL$, momentum scale $P_0 = M\sqrt{gL}$, and time scale $\tau_0 = \sqrt{L/g}$. Show that the dimensionless Hamiltonian becomes

$$\bar{H} = \tfrac{1}{2}\bar{P}^2 + \bar{X} + \frac{\bar{p}^2}{2r} + r\bar{x} \ ,$$

with $r = m/M$, and with equations of motion $dX/dt = \partial\bar{H}/\partial\bar{P}$, *etc.* (Here the bar indicates dimensionless variables: $\bar{P} = P/P_0$, $\bar{t} = t/\tau_0$, *etc.*) What special dynamical consequences hold for $r = 1$?

(b) Compute the microcanonical average piston height $\langle X \rangle$. The analogous dynamical average is

$$\langle X \rangle_T = \lim_{T\to\infty} \frac{1}{T} \int\limits_0^T dt\, X(t) \ .$$

When computing microcanonical averages, it is helpful to use the Laplace transform, discussed toward the end of §3.3 of the notes. (It is possible to compute the microcanonical average by more brute force methods as well.)

(c) Compute the microcanonical average of the rate of collisions between the ball and the floor. Show that this is given by

$$\Big\langle \sum_i \delta(t - t_i) \Big\rangle = \big\langle \Theta(v)\, v\, \delta(x - 0^+) \big\rangle \ .$$

The analogous dynamical average is

$$\langle \gamma \rangle_T = \lim_{T\to\infty} \frac{1}{T} \int\limits_0^T dt \sum_i \delta(t - t_i) \ ,$$

where $\{t_i\}$ is the set of times at which the ball hits the floor.

(d) How do your results change if you do not enforce the dynamical constraint $X \geq x$?

(e) Write a computer program to simulate this system. The only input should be the mass ratio r (set $\bar{E} = 10$ to fix the energy). You also may wish to input the initial conditions, or perhaps to choose the initial conditions randomly (all satisfying energy conservation, of course!). Have your program compute the microcanonical as well as dynamical averages in parts (b) and (c). Plot out the Poincaré section of P *vs.* X for those times when the ball hits the floor. Investigate this for several values of r. Just to show you that this is interesting, I've plotted some of my own numerical results in fig. 3.

Solution:

(a) Once we choose a length scale L (arbitrary), we may define $E_0 = M\,gL$, $P_0 = M\sqrt{gL}$, $V_0 = \sqrt{gL}$, and $\tau_0 = \sqrt{L/g}$ as energy, momentum, velocity, and time scales, respectively,

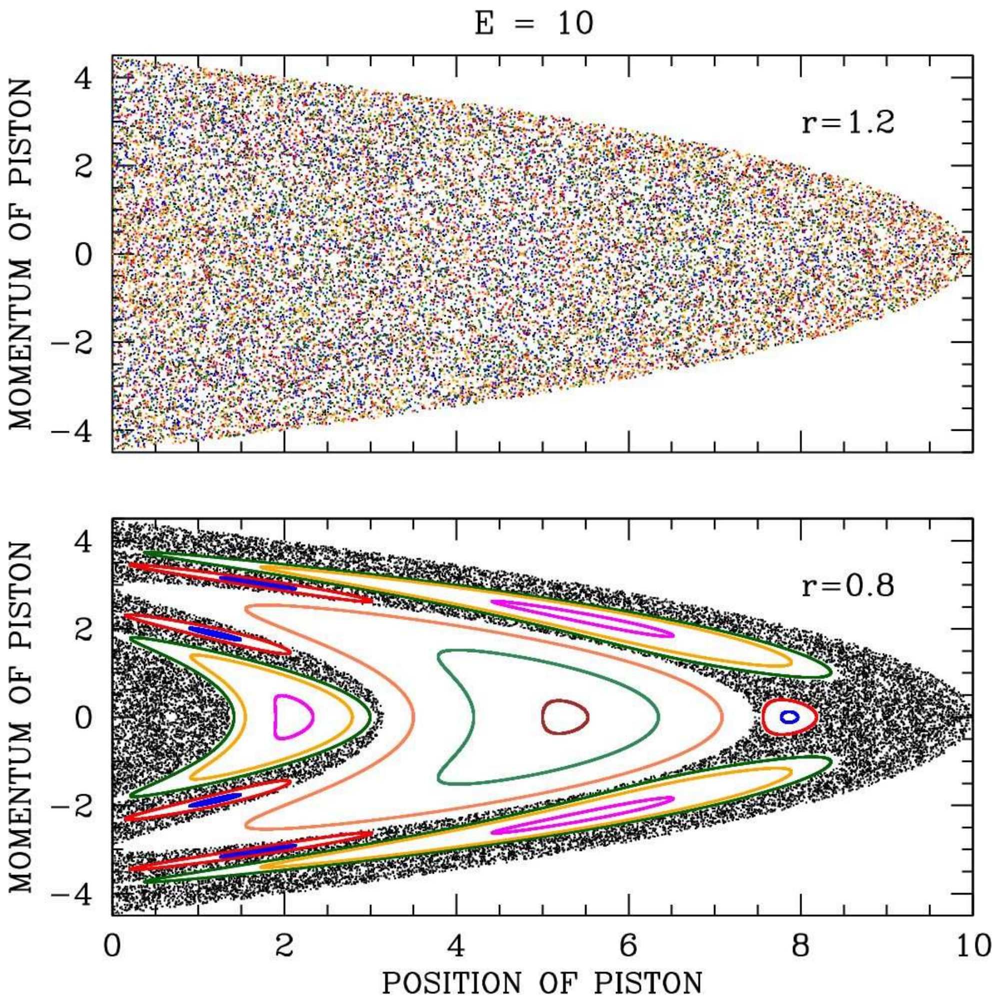

Figure 3: Poincaré sections for the ball and piston head problem. Each color corresponds to a different initial condition. When the mass ratio $r = m/M$ exceeds unity, the system apparently becomes ergodic.

the result follows directly. Rather than write $\bar{P} = P/P_0$ *etc.*, we will drop the bar notation and write

$$H = \tfrac{1}{2}P^2 + X + \frac{p^2}{2r} + rx \ .$$

(b) What is missing from the Hamiltonian of course is the interaction potential between the ball and the piston head. We assume that both objects are impenetrable, so the potential energy is infinite when the two overlap. We further assume that the ball is a point particle (otherwise reset ground level to minus the diameter of the ball). We can eliminate the interaction potential from H if we enforce that each time $X = x$ the ball and the piston head undergo an elastic collision. From energy and momentum conservation, it is easy to derive the elastic collision formulae

$$P' = \frac{1-r}{1+r}\; P + \frac{2}{1+r}\; p$$

$$p' = \frac{2r}{1+r}\; P - \frac{1-r}{1+r}\; p \; .$$

We can now answer the last question from part (a). When $r = 1$, we have that $P' = p$ and $p' = P$, *i.e.* the ball and piston simply exchange momenta. The problem is then equivalent to two identical particles elastically bouncing off the bottom of the piston, and moving through each other as if they were completely transparent. When the trajectories cross, however, the particles exchange identities.

Averages within the microcanonical ensemble are normally performed with respect to the phase space distribution

$$\varrho(\boldsymbol{\varphi}) = \frac{\delta\big((E - H(\boldsymbol{\varphi})\big)}{\mathsf{Tr}\;\delta\big(E - H(\boldsymbol{\varphi})\big)} \; ,$$

where $\boldsymbol{\varphi} = (P, X, p, x)$, and

$$\mathsf{Tr}\; F(\boldsymbol{\varphi}) = \int\limits_{-\infty}^{\infty}\!dP \int\limits_{0}^{\infty}\!dX \int\limits_{-\infty}^{\infty}\!dp \int\limits_{0}^{\infty}\!dx\; F(P, X, p, x) \; .$$

Since $X \geq x$ is a dynamical constraint, we should define an appropriately restricted microcanonical average:

$$\big\langle F(\boldsymbol{\varphi})\big\rangle_{\mu\mathrm{ce}} \equiv \widetilde{\mathsf{Tr}}\left[F(\boldsymbol{\varphi})\,\delta\big(E - H(\boldsymbol{\varphi})\big)\right] \Big/ \widetilde{\mathsf{Tr}}\,\delta\big(E - H(\boldsymbol{\varphi})\big)$$

where

$$\widetilde{\mathsf{Tr}}\,F(\boldsymbol{\varphi}) \equiv \int\limits_{-\infty}^{\infty}\!dP \int\limits_{0}^{\infty}\!dX \int\limits_{-\infty}^{\infty}\!dp \int\limits_{0}^{X}\!dx\; F(P, X, p, x)$$

is the modified trace. Note that the integral over x has an upper limit of X rather than ∞, since the region of phase space with $x > X$ is *dynamically inaccessible.*

When computing the traces, we shall make use of the following result from the theory of Laplace transforms. The Laplace transform of a function $K(E)$ is

$$\widehat{K}(\beta) = \int\limits_{0}^{\infty}\!dE\; K(E)\, e^{-\beta E} \; .$$

The inverse Laplace transform is given by

$$K(E) = \int\limits_{c-i\infty}^{c+i\infty} \frac{d\beta}{2\pi i}\; \widehat{K}(\beta)\, e^{\beta E} \; ,$$

where the integration contour, which is a line extending from $\beta = c - i\infty$ to $\beta = c + i\infty$, lies to the right of any singularities of $\widehat{K}(\beta)$ in the complex β-plane. For this problem, all we shall need is the following:

$$K(E) = \frac{E^{t-1}}{\Gamma(t)} \qquad \Longleftrightarrow \qquad \widehat{K}(\beta) = \beta^{-t} \ .$$

For a proof, see §3.3.1 of the lecture notes.

We're now ready to compute the microcanonical average of X. We have

$$\langle X \rangle = \frac{N(E)}{D(E)} \ ,$$

where

$$\begin{aligned} N(E) &= \widetilde{\mathsf{Tr}} \left[X \, \delta(E - H) \right] \\ D(E) &= \widetilde{\mathsf{Tr}} \, \delta(E - H) \ . \end{aligned}$$

Let's first compute $D(E)$. To do this, we compute the Laplace transform $\widehat{D}(\beta)$:

$$\begin{aligned} \widehat{D}(\beta) &= \widetilde{\mathsf{Tr}} \, e^{-\beta H} \\ &= \int\limits_{-\infty}^{\infty} dP \, e^{-\beta P^2/2} \int\limits_{-\infty}^{\infty} dp \, e^{-\beta p^2/2r} \int\limits_{0}^{\infty} dX \, e^{-\beta X} \int\limits_{0}^{X} dx \, e^{-\beta r x} \\ &= \frac{2\pi\sqrt{r}}{\beta} \int\limits_{0}^{\infty} dX \, e^{-\beta X} \left(\frac{1 - e^{-\beta r X}}{\beta r} \right) = \frac{\sqrt{r}}{1+r} \cdot \frac{2\pi}{\beta^3} \ . \end{aligned}$$

Similarly for $\widehat{N}(\beta)$ we have

$$\begin{aligned} \widehat{N}(\beta) &= \widetilde{\mathsf{Tr}} \, X \, e^{-\beta H} \\ &= \int\limits_{-\infty}^{\infty} dP \, e^{-\beta P^2/2} \int\limits_{-\infty}^{\infty} dp \, e^{-\beta p^2/2r} \int\limits_{0}^{\infty} dX \, X \, e^{-\beta X} \int\limits_{0}^{X} dx \, e^{-\beta r x} \\ &= \frac{2\pi\sqrt{r}}{\beta} \int\limits_{0}^{\infty} dX \, X \, e^{-\beta X} \left(\frac{1 - e^{-\beta r X}}{\beta r} \right) = \frac{(2+r)\, r^{3/2}}{(1+r)^2} \cdot \frac{2\pi}{\beta^4} \ . \end{aligned}$$

Taking the inverse Laplace transform, we then have

$$D(E) = \frac{\sqrt{r}}{1+r} \cdot \pi E^2 \qquad , \qquad N(E) = \frac{(2+r)\,\sqrt{r}}{(1+r)^2} \cdot \tfrac{1}{3} \pi E^3 \ .$$

We then have

$$\langle X \rangle = \frac{N(E)}{D(E)} = \left(\frac{2+r}{1+r} \right) \cdot \tfrac{1}{3} E \ .$$

The ‘brute force’ evaluation of the integrals isn’t so bad either. We have

$$D(E) = \int\limits_{-\infty}^{\infty}\!dP\int\limits_{0}^{\infty}\!dX\int\limits_{-\infty}^{\infty}\!dp\int\limits_{0}^{X}\!dx\;\delta\big(\tfrac{1}{2}P^2 + \tfrac{1}{2r}p^2 + X + rx - E\big)\ .$$

To evaluate, define $P = \sqrt{2}\,u_x$ and $p = \sqrt{2r}\,u_y$. Then we have $dP\,dp = 2\sqrt{r}\,du_x\,du_y$ and $\frac{1}{2}P^2 + \frac{1}{2r}\,p^2 = u_x^2 + u_y^2$. Now convert to 2D polar coordinates with $w \equiv u_x^2 + u_y^2$. Thus,

$$\begin{aligned}
D(E) &= 2\pi\sqrt{r}\int\limits_{0}^{\infty}\!dw\int\limits_{0}^{\infty}\!dX\int\limits_{0}^{X}\!dx\;\delta\big(w + X + rx - E\big)\\
&= \frac{2\pi}{\sqrt{r}}\int\limits_{0}^{\infty}\!dw\int\limits_{0}^{\infty}\!dX\int\limits_{0}^{X}\!dx\;\;\Theta(E-w-X)\,\Theta(X+rX-E+w)\\
&= \frac{2\pi}{\sqrt{r}}\int\limits_{0}^{E}\!dw\int\limits_{\frac{E-w}{1+r}}^{E-w}\!dX = \frac{2\pi\sqrt{r}}{1+r}\int\limits_{0}^{E}\!dq\,q = \frac{\sqrt{r}}{1+r}\cdot\pi E^2\ ,
\end{aligned}$$

with $q = E - w$. Similarly,

$$\begin{aligned}
N(E) &= 2\pi\sqrt{r}\int\limits_{0}^{\infty}\!dw\int\limits_{0}^{\infty}\!dX\;X\int\limits_{0}^{X}\!dx\;\delta\big(w + X + rx - E\big)\\
&= \frac{2\pi}{\sqrt{r}}\int\limits_{0}^{\infty}\!dw\int\limits_{0}^{\infty}\!dX\;X\int\limits_{0}^{X}\!dx\;\;\Theta(E-w-X)\,\Theta(X+rX-E+w)\\
&= \frac{2\pi}{\sqrt{r}}\int\limits_{0}^{E}\!dw\int\limits_{\frac{E-w}{1+r}}^{E-w}\!dX\;X = \frac{2\pi}{\sqrt{r}}\int\limits_{0}^{E}\!dq\left(1-\frac{1}{(1+r)^2}\right)\cdot\tfrac{1}{2}q^2 = \left(\frac{2+r}{1+r}\right)\cdot\frac{\sqrt{r}}{1+r}\cdot\tfrac{1}{3}\pi E^3\ .
\end{aligned}$$

(c) Using the general result

$$\delta\big(F(x) - A\big) = \sum_i \frac{\delta(x - x_i)}{\big|F'(x_i)\big|}\ ,$$

where $F(x_i) = A$, we recover the desired expression. We should be careful not to double count, so to avoid this difficulty we can evaluate $\delta(t - t_i^+)$, where $t_i^+ = t_i + 0^+$ is infinitesimally later than t_i. The point here is that when $t = t_i^+$ we have $p = r\,v > 0$ (*i.e.* just after hitting the bottom). Similarly, at times $t = t_i^-$ we have $p < 0$ (*i.e.* just prior to hitting the bottom). Note $v = p/r$. Again we write $\gamma(E) = N(E)/D(E)$, this time with

$$N(E) = \widetilde{\mathsf{Tr}}\,\big[\Theta(p)\,r^{-1}p\,\delta(x - 0^+)\,\delta(E - H)\big]\ .$$

The Laplace transform is

$$
\begin{aligned}
\widehat{N}(\beta) &= \int\limits_{-\infty}^{\infty} dP\, e^{-\beta P^2/2} \int\limits_0^{\infty} dp\; r^{-1}\, p\, e^{-\beta p^2/2r} \int\limits_0^{\infty} dX\, e^{-\beta X} \\
&= \sqrt{\frac{2\pi}{\beta}} \cdot \frac{1}{\beta} \cdot \frac{1}{\beta} = \sqrt{2\pi}\, \beta^{-5/2} \ .
\end{aligned}
$$

Thus,

$$
N(E) = \tfrac{4\sqrt{2}}{3}\, E^{3/2}
$$

and

$$
\langle \gamma \rangle = \frac{N(E)}{D(E)} = \tfrac{4\sqrt{2}}{3\pi} \left(\frac{1+r}{\sqrt{r}} \right) E^{-1/2} \ .
$$

(d) When the constraint $X \geq x$ is removed, we integrate over all phase space. We then have

$$
\begin{aligned}
\widehat{D}(\beta) &= \mathsf{Tr}\; e^{-\beta H} \\
&= \int\limits_{-\infty}^{\infty} dP\, e^{-\beta P^2/2} \int\limits_{-\infty}^{\infty} dp\, e^{-\beta p^2/2r} \int\limits_0^{\infty} dX\, e^{-\beta X} \int\limits_0^{\infty} dx\, e^{-\beta r x} = \frac{2\pi\sqrt{r}}{\beta^3} \ .
\end{aligned}
$$

For part (b) we would then have

$$
\begin{aligned}
\widehat{N}(\beta) &= \mathsf{Tr}\; X\, e^{-\beta H} \\
&= \int\limits_{-\infty}^{\infty} dP\, e^{-\beta P^2/2} \int\limits_{-\infty}^{\infty} dp\, e^{-\beta p^2/2r} \int\limits_0^{\infty} dX\, X\, e^{-\beta X} \int\limits_0^{\infty} dx\, e^{-\beta r x} = \frac{2\pi\sqrt{r}}{\beta^4} \ .
\end{aligned}
$$

The respective inverse Laplace transforms are $D(E) = \pi\sqrt{r}\, E^2$ and $N(E) = \frac{1}{3}\pi\sqrt{r}\, E^3$. The microcanonical average of X would then be

$$
\langle X \rangle = \tfrac{1}{3} E \ .
$$

Using the restricted phase space, we obtained a value which is greater than this by a factor of $(2+r)/(1+r)$. That the restricted average gives a larger value makes good sense, since X is not allowed to descend below x in that case. For part (c), we would obtain the same result for $N(E)$ since $x = 0$ in the average. We would then obtain

$$
\langle \gamma \rangle = \tfrac{4\sqrt{2}}{3\pi}\, r^{-1/2}\, E^{-1/2} \ .
$$

The restricted microcanonical average yields a rate which is larger by a factor $1+r$. Again, it makes good sense that the restricted average should yield a higher rate, since the ball is not allowed to attain a height greater than the instantaneous value of X.

(e) It is straightforward to simulate the dynamics. So long as $0 < x(t) < X(t)$, we have

$$
\dot{X} = P \quad , \quad \dot{P} = -1 \quad , \quad \dot{x} = \frac{p}{r} \quad , \quad \dot{p} = -r \ .
$$

Starting at an arbitrary time t_0, these equations are integrated to yield

$$\begin{aligned} X(t) &= X(t_0) + P(t_0)\,(t-t_0) - \tfrac{1}{2}(t-t_0)^2 \\ P(t) &= P(t_0) - (t-t_0) \\ x(t) &= x(t_0) + \frac{p(t_0)}{r}\,(t-t_0) - \tfrac{1}{2}(t-t_0)^2 \\ p(t) &= p(t_0) - r(t-t_0)\ . \end{aligned}$$

We must stop the evolution when one of two things happens. The first possibility is a bounce at $t = t_\mathrm{b}$, meaning $x(t_\mathrm{b}) = 0$. The momentum $p(t)$ changes discontinuously at the bounce, with $p(t_\mathrm{b}^+) = -p(t_\mathrm{b}^-)$, and where $p(t_\mathrm{b}^-) < 0$ necessarily. The second possibility is a collision at $t = t_\mathrm{c}$, meaning $X(t_\mathrm{c}) = x(t_\mathrm{c})$. Integrating across the collision, we must conserve both energy and momentum. This means

$$P(t_\mathrm{c}^+) = \frac{1-r}{1+r}\,P(t_\mathrm{c}^-) + \frac{2}{1+r}\,p(t_\mathrm{c}^-)$$

$$p(t_\mathrm{c}^+) = \frac{2r}{1+r}\,P(t_\mathrm{c}^-) - \frac{1-r}{1+r}\,p(t_\mathrm{c}^-)\ .$$

In the following tables I report on the results of numerical simulations, comparing dynamical averages with (restricted) phase space averages within the microcanonical ensemble. For $r = 0.3$ the microcanonical averages poorly approximate the dynamical averages, and the dynamical averages are dependent on the initial conditions, indicating that the system is not ergodic. For $r = 1.2$, the agreement between dynamical and microcanonical averages generally improves with averaging time. Indeed, it has been shown by N. I. Chernov, *Physica D* **53**, 233 (1991), building on the work of M. P. Wojtkowski, *Comm. Math. Phys.* **126**, 507 (1990) that this system is ergodic for $r > 1$. Wojtkowski also showed that this system is equivalent to the *wedge billiard*, in which a single point particle of mass m bounces inside a two-dimensional wedge-shaped region $\big\{(\mathsf{x},\mathsf{y})\,\big|\,\mathsf{x} \ge 0\,,\ \mathsf{y} \ge \mathsf{x}\,\mathrm{ctn}\,\phi\big\}$ for some fixed angle $\phi = \tan^{-1}\sqrt{\frac{m}{M}}$. To see this, pass to relative ($\mathcal{X}$) and center-of-mass ($\mathcal{Y}$) coordinates,

$$\mathcal{X} = X - x \qquad\qquad \mathcal{P}_x = \frac{mP - Mp}{M+m}$$

$$\mathcal{Y} = \frac{MX + mx}{M+m} \qquad\qquad \mathcal{P}_y = P + p\ .$$

Then

$$H = \frac{(M+m)\,\mathcal{P}_x^2}{2Mm} + \frac{\mathcal{P}_y^2}{2(M+m)} + (M+m)\,g\mathcal{Y}\ .$$

There are two constraints. One requires $X \ge x$, *i.e.* $\mathcal{X} \ge 0$. The second requires $x > 0$, *i.e.*

$$x = \mathcal{Y} - \frac{M}{M+m}\,\mathcal{X} \ge 0\ .$$

Now define $\mathsf{x} \equiv \mathcal{X}$, $\mathsf{p}_x \equiv \mathcal{P}_x$, and rescale $\mathsf{y} \equiv \frac{M+m}{\sqrt{Mm}}\,\mathcal{Y}$ and $\mathsf{p}_y \equiv \frac{\sqrt{Mm}}{M+m}\,\mathcal{P}_y$ to obtain

$$H = \frac{1}{2\mu}\big(\mathsf{p}_x^2 + \mathsf{p}_y^2\big) + \mathsf{M}\,g\,\mathsf{y}$$

with $\mu = \frac{Mm}{M+m}$ the familiar reduced mass and $\mathsf{M} = \sqrt{Mm}$. The constraints are then $\mathsf{x} \geq 0$ and $\mathsf{y} \geq \sqrt{\frac{M}{m}}\,\mathsf{x}$.

r	$X(0)$	$\langle X(t)\rangle$	$\langle X\rangle_{\mu\text{ce}}$	$\langle\gamma(t)\rangle$	$\langle\gamma\rangle_{\mu\text{ce}}$
0.3	0.1	6.1743	5.8974	0.5283	0.4505
0.3	1.0	5.7303	5.8974	0.4170	0.4505
0.3	3.0	5.7876	5.8974	0.4217	0.4505
0.3	5.0	5.8231	5.8974	0.4228	0.4505
0.3	7.0	5.8227	5.8974	0.4228	0.4505
0.3	9.0	5.8016	5.8974	0.4234	0.4505
0.3	9.9	6.1539	5.8974	0.5249	0.4505

Table 2: Comparison of time averages and microcanonical ensemble averages for $r = 0.3$. Initial conditions are $P(0) = x(0) = 0$, with $X(0)$ given in the table and $E = 10$. Averages were performed over a period extending for $N_{\rm b} = 10^7$ bounces.

r	$X(0)$	$\langle X(t)\rangle$	$\langle X\rangle_{\mu\text{ce}}$	$\langle\gamma(t)\rangle$	$\langle\gamma\rangle_{\mu\text{ce}}$
1.2	0.1	4.8509	4.8545	0.3816	0.3812
1.2	1.0	4.8479	4.8545	0.3811	0.3812
1.2	3.0	4.8493	4.8545	0.3813	0.3812
1.2	5.0	4.8482	4.8545	0.3813	0.3812
1.2	7.0	4.8472	4.8545	0.3808	0.3812
1.2	9.0	4.8466	4.8545	0.3808	0.3812
1.2	9.9	4.8444	4.8545	0.3807	0.3812

Table 3: Comparison of time averages and microcanonical ensemble averages for $r = 1.2$. Initial conditions are $P(0) = x(0) = 0$, with $X(0)$ given in the table and $E = 10$. Averages were performed over a period extending for $N_{\rm b} = 10^7$ bounces.

Finally, in fig. 4, I plot the running averages of $X_{\rm av}(t) \equiv t^{-1}\int\limits_0^t dt'\,X(t')$ for the cases $r = 0.3$ and $r = 1.2$, each with $E = 10$, and each for three different sets of initial conditions. For $r = 0.3$, the system is not ergodic, and the dynamics will be restricted to a subset of phase space. Accordingly the long time averages vary with the initial conditions. For $r = 1.2$ the system is ergodic and the results converge to the appropriate restricted microcanonical average $\langle X\rangle_{\mu\text{ce}}$ at large times, independent of initial conditions.

r	$X(0)$	N_b	$\langle X(t)\rangle$	$\langle X\rangle_{\mu\mathrm{ce}}$	$\langle\gamma(t)\rangle$	$\langle\gamma\rangle_{\mu\mathrm{ce}}$
1.2	7.0	10^4	4.8054892	4.8484848	0.37560388	0.38118510
1.2	7.0	10^5	4.8436969	4.8484848	0.38120356	0.38118510
1.2	7.0	10^6	4.8479414	4.8484848	0.38122778	0.38118510
1.2	7.0	10^7	4.8471686	4.8484848	0.38083749	0.38118510
1.2	7.0	10^8	4.8485825	4.8484848	0.38116282	0.38118510
1.2	7.0	10^9	4.8486682	4.8484848	0.38120259	0.38118510
1.2	1.0	10^9	4.8485381	4.8484848	0.38118069	0.38118510
1.2	9.9	10^9	4.8484886	4.8484848	0.38116295	0.38118510

Table 4: Comparison of time averages and microcanonical ensemble averages for $r = 1.2$, with N_b ranging from 10^4 to 10^9.

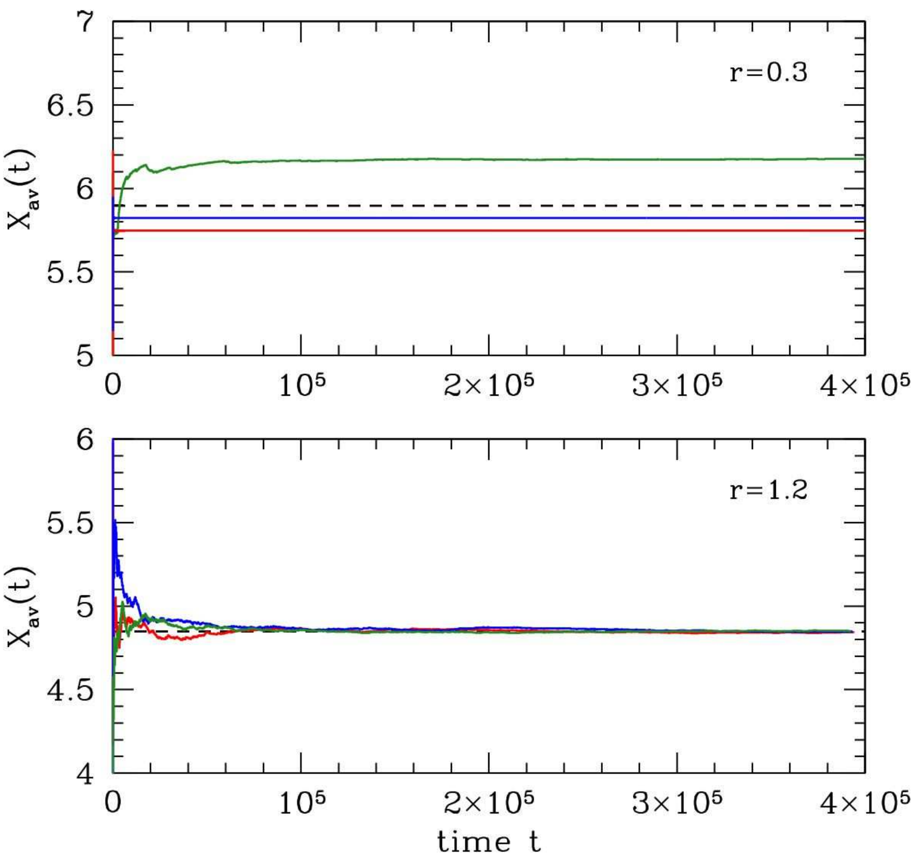

Figure 4: Long time running numerical averages $X_\mathrm{av}(t) \equiv t^{-1}\int_0^t dt'\, X(t')$ for $r = 0.3$ (top) and $r = 1.2$ (bottom), each for three different initial conditions, with $E = 10$ in all cases. Note how in the $r = 0.3$ case the long time average is dependent on the initial condition, while the $r = 1.2$ case is ergodic and hence independent of initial conditions. The dashed black line shows the restricted microcanonical average, $\langle X\rangle_{\mu\mathrm{ce}} = \frac{(2+r)}{(1+r)} \cdot \frac{1}{3}E$.