Name: **Huan Q. Bui**
Course: **8.333 - Statistical Mechanics I**
Problem set: **#2**

1. Random deposition.

(a) Consider a site. Assume that the gold atoms arrive at this site over time via a Poisson process. The deposition rate is d layers per second, so this Poisson process has rate d. Over time t, the average number of deposition at a site is dt. With this, we have

$$\Pr(m \text{ atoms in time } t) = \boxed{\frac{(dt)^m e^{-dt}}{m!}}$$

Glass is not covered if there is no deposition, i.e., $m = 0$. The fraction of the glass not covered by the atoms is the probability of zero deposition:

$$\Pr(0 \text{ atoms in time } t) = \boxed{e^{-dt}}$$

We see that the fraction of the glass not covered decreases exponentially in time.

(b) From Part (a), we know that the average thickness is $\langle x \rangle = dt$. To find the variance we need to compute the second moment:

$$\langle x^2 \rangle = \sum_{i=0}^{\infty} x^2 \frac{(dt)^x e^{-dt}}{x!} = dt + d^2t^2.$$

Therefore, the variance in thickness is

$$\sigma^2 = \langle x^2 \rangle - \langle x \rangle^2 = \boxed{dt}$$

2. Semi-flexible polymer in two dimensions.

(a) It's nicer to work with the ϕ-dependent $\mathcal{H}$, so let us write $\mathbf{t}_m \cdot \mathbf{t}_n$ in terms of angles:

$$\mathbf{t}_m \cdot \mathbf{t}_n = a^2 \cos(\theta_m + \theta_{m+1} + \cdots + \theta_{n-1}).$$

The summed form for the angles is not very convenient to work with, as there is no clear way to find $\langle \mathbf{t}_m \cdot \mathbf{t}_n \rangle$ written in this form. Instead, let us find $a^2\langle \exp(i(\theta_m + \theta_{m+1} + \cdots + \theta_{n-1}))\rangle$ and then take the real part. This, written in this form, is still cumbersome. However, we may assume that the angles ϕ_i are independent, and therefore

$$\langle \mathbf{t}_m \cdot \mathbf{t}_n \rangle = \text{Re}\left[a^2 \langle \exp(i(\theta_m + \theta_{m+1} + \cdots + \theta_{n-1}))\rangle\right] = \text{Re}\left[a^2 \prod_{j=m}^{n-1} \langle e^{i\phi_j}\rangle\right] = a^2 \prod_{j=m}^{n-1} \langle \cos \phi_j \rangle.$$

Moreover, since the angles ϕ_i's are independent, the probability for each configuration is simply the product of the individual probabilities:

$$\Pr(\phi_1, \dots, \phi_{N-1}) = \exp\left[\frac{a^2\kappa}{k_BT} \sum_{i=1}^{N-1} \cos\phi_i\right] = \prod_{i=1}^{N-1} \exp\left[\frac{a^2\kappa}{k_BT} \cos\phi_i\right].$$

And so we may write

$$\Pr(\phi_i) = \exp\left[\frac{a^2\kappa}{k_BT} \cos\phi_i\right]$$

With this we have

$$\langle \mathbf{t}_m \cdot \mathbf{t}_n \rangle = a^2 \prod_{j=m}^{n-1} \frac{\int d\phi \cos\phi \exp\left[\frac{a^2\kappa}{k_B T}\cos\phi_i\right]}{\int d\phi \exp\left[\frac{a^2\kappa}{k_B T}\cos\phi_i\right]} = a^2 \left\{ \frac{\int d\phi \cos\phi \exp\left[\frac{a^2\kappa}{k_B T}\cos\phi\right]}{\int d\phi \exp\left[\frac{a^2\kappa}{k_B T}\cos\phi\right]} \right\}^{|n-m|}.$$

So $\langle \mathbf{t}_m \cdot \mathbf{t}_n \rangle$ has the form $a^2[f(T)]^{|n-m|}$ where $f(T)$ is the fraction in the curly brackets. We may write $\langle \mathbf{t}_m \cdot \mathbf{t}_n \rangle$ as an exponential:

$$\langle \mathbf{t}_m \cdot \mathbf{t}_n \rangle = a^2 \exp\left[|n-m| \ln f(T)\right] = a^2 \exp\left[\frac{|n-m|}{1/\ln f(T)}\right] \equiv a^2 \exp\left[\frac{-|n-m|}{\xi}\right],$$

as desired. The persistence length is thus

$$\boxed{l_p = a\xi = \frac{a}{-\ln f(T)} = \frac{a}{\ln\left[\frac{\int d\phi \exp\left[\frac{a^2\kappa}{k_B T}\cos\phi\right]}{\int d\phi \cos\phi \exp\left[\frac{a^2\kappa}{k_B T}\cos\phi\right]}\right]}}$$

(b) By definition, we have

$$\mathbf{R} = \sum_{i=1}^{N} \mathbf{t}_i \implies \langle R^2 \rangle = \langle \mathbf{R} \cdot \mathbf{R} \rangle = \sum_{m,n=1}^{N} \langle \mathbf{t}_m, \mathbf{t}_n \rangle = \sum_{m,n=1}^{N} a^2 \exp\left[\frac{-|n-m|}{\xi}\right].$$

Now we consider what happens when $N \to \infty$. We see that $\mathbf{R}$ has the form

$$\langle R^2 \rangle = a^2 \left[N + N_1 e^{-1/\xi} + N_2 e^{-2/\xi} + N_3 e^{-3/\xi} + \dots\right]$$

where $N_1, N_2, N_3, \dots$ are natural numbers. In the limit $N \to \infty$, we have $N_j \approx 2N$ for small j's (swapping n, m gives an extra factor of 2), and N_j's for large j's don't really matter because of the exponential decay $e^{-j/\xi}$. So, we may very well write this as

$$\langle R^2 \rangle \approx a^2 \left[N + 2N\left(e^{-1/\xi} + e^{-2/\xi} + e^{-3/\xi} + \dots\right)\right]$$

We now recall that

$$\frac{x}{1-x} = x + x^2 + x^3 + \dots$$

So, we have a rather compact formula for $\langle R^2 \rangle$:

$$\boxed{\langle R^2 \rangle = a^2 N \left(1 + 2\frac{e^{-1/\xi}}{1 - e^{-1/\xi}}\right), \qquad N \to \infty}$$

(c) $\mathbf{R}$ is a sum of iid's $\mathbf{t}_i$. In view of the central limit theorem, $p(\mathbf{R})$ is a Gaussian. To determine the form of $p(\mathbf{R})$, we must find the first and second moments. Since each $\mathbf{t}_i$ is random, we can conclude that $\langle \mathbf{R} \rangle = 0$. The second moment is given by Part (b), and so the variance of this distribution is $\sigma^2 = \langle R^2 \rangle - 0 = \langle R^2 \rangle$ which is what we found in Part (b). To find the normalization constant, we look at the covariance matrix C. Its determinant $|\det(C)|$ will be the product of $\langle R_x^2 \rangle$ and $\langle R_y^2 \rangle$, each of which is $\langle R^2 \rangle/2$ (by symmetry, and the fact that variances of independent variables add). With these,

$$p(\mathbf{R}) = \frac{1}{\sqrt{(2\pi)^2 |\det(C)|}} \exp\left(-\frac{\mathbf{R}^\top C^{-1} \mathbf{R}}{2}\right) = \boxed{\frac{1}{\pi \langle R^2 \rangle} \exp\left(-\frac{\mathbf{R} \cdot \mathbf{R}}{\langle R^2 \rangle}\right)}$$

where we have used the fact that $C = \langle R^2 \rangle \mathbb{I}/2$.

(d) We shall "formally" consider the modified probability weight:

$$\exp(-\mathcal{H}/k_BT) \to \exp(\mathbf{F}\cdot\mathbf{R}/k_BT)\exp(-\mathcal{H}/k_BT).$$

Taking the average of $\mathbf{R}$ under these new weights yields

$$\langle \mathbf{R}\rangle = \frac{\int \mathbf{R}\exp(\mathbf{F}\cdot\mathbf{R}/k_BT)\exp(-\mathcal{H}/k_BT)}{\int \exp(\mathbf{F}\cdot\mathbf{R}/k_BT)\exp(-\mathcal{H}/k_BT)}.$$

We may treat $\mathbf{R}\exp(\mathbf{F}\cdot\mathbf{R}/k_BT)$ and $\exp(\mathbf{F}\cdot\mathbf{R}/k_BT)$ as input functions whose averages we wish to find and write

$$\langle R\rangle = \frac{\langle \mathbf{R}\exp(\mathbf{F}\cdot\mathbf{R}/k_BT)\rangle'}{\langle \exp(\mathbf{F}\cdot\mathbf{R}/k_BT)\rangle'}$$

where the new average $\langle\cdot\rangle'$ are essentially averages for when $\mathbf{F}=0$. The denominator becomes unity, while the numerator can be expanded as

$$\langle \mathbf{R}_i\exp(\mathbf{F}\cdot\mathbf{R}/k_BT)\rangle' \approx \langle \mathbf{R}_i\rangle' + \frac{\langle \mathbf{R}_i\mathbf{F}\cdot\mathbf{R}\rangle'}{k_BT} + \frac{\langle \mathbf{R}_i(\mathbf{F}\cdot\mathbf{R})^2\rangle'}{2k_BT} + \frac{\langle \mathbf{R}_i(\mathbf{F}\cdot\mathbf{R})^3\rangle'}{6k_BT} + \dots$$

where (1) $\mathbf{R}_i$'s are the components of $\mathbf{R}$ (i.e., i is x and y), and that $\langle R\rangle$ at $\mathbf{F}=0$ is simply $\langle R^2\rangle$ which we already know. Moreover, terms with odd-powered $\mathbf{R}$'s will vanish by symmetry. We are thus interested in the term

$$\langle \mathbf{R}_i\mathbf{F}\cdot\mathbf{R}\rangle' = \langle \mathbf{R}_i\mathbf{F}_j\mathbf{R}_j\rangle' = \mathbf{F}_j\langle \mathbf{R}_i\mathbf{R}_j\rangle' = F_j\delta_{ij}\langle R^2\rangle/2 = F_i\langle R^2\rangle/2.$$

With this we have

$$\boxed{\langle \mathbf{R}\rangle = \frac{\langle R^2\rangle}{2k_BT}\mathbf{F} + \mathcal{O}(F^3) = K^{-1}\mathbf{F} + \mathcal{O}(F^3), \qquad K = \frac{2k_BT}{\langle R^2\rangle}}$$

3. Foraging. We have

$$p(r|t) = \frac{r}{2Dt}\exp\left(-\frac{r^2}{4Dt}\right) \quad \text{and} \quad p(t) \propto \exp\left(-\frac{t}{\tau}\right).$$

Normalizing $p(t)$ give the leading factor equal $1/\tau$. So, we can write

$$p(t) = \frac{1}{\tau}\exp\left(-\frac{t}{\tau}\right).$$

The (unconditional) probability of finding the searcher at a distance r from the nest is

$$p(r) = \int_0^\infty p(r|t)p(t)\,dt = \int_0^\infty \frac{r}{2D\tau t}\exp\left(-\frac{r^2}{4Dt} - \frac{t}{\tau}\right)\,dt.$$

We may compute this using saddle-point approximation. Let $f(t) = r^2/4Dt + t/\tau$. Then we see that f attains a maximum at $t_0 = r\sqrt{\tau}/2\sqrt{D}$. Let $g(t) = r/2D\tau t$. The saddle point approximation reads

$$p(r) \approx e^{-f(t_0)}g(t_0)\sqrt{\frac{2\pi}{f''(t_0)}} = \exp\left(\frac{-r}{\sqrt{D\tau}}\right)\frac{1}{\sqrt{D}\tau^{3/2}}\sqrt{\frac{2\pi r t^{3/2}}{4\sqrt{D}}} \implies \boxed{p(r) \propto \sqrt{r}\exp\left(\frac{-r}{\sqrt{D\tau}}\right)}$$

We could also say that asymptotically the exponential decay in r dominates over the $\sqrt{r}$ growth, and so in the large r limit, $p(r) \sim \exp\left(-r/\sqrt{D\tau}\right)$.

4. Jensen's inequality and Kullback–Liebler divergence.

(a) <u>Claim:</u> Jensen's inequality: For a convex function $\langle f(x)\rangle \geq f(\langle x\rangle)$. To prove this, we let a probability density function $p(x)$ be given. Then

$$\langle f(x)\rangle = \int p(x)f(x)\,dx \geq \int p(x)\left[f(\langle x\rangle) + f'(\langle x\rangle)(x-\langle x\rangle)\right]\,dx = f(\langle x\rangle) + f'(\langle x\rangle)\langle x - \langle x\rangle\rangle = f(\langle x\rangle).$$

And we're done.

(b) We see that $D(p|q)$ is an expectation taken with respect to $p(x)$. The proof uses Jensen's inequality with the fact that the function $-\ln(x)$ is convex:

$$\begin{aligned} D(p|q) &= \left\langle \ln\frac{p}{q}\right\rangle_p \\ &= \left\langle -\ln\frac{q}{p}\right\rangle_p \\ &\geq -\ln\left\langle\frac{q}{p}\right\rangle_p \\ &= -\ln\left[\int p(x)\frac{q(x)}{p(x)}\,dx\right] = 0. \end{aligned}$$

5. The book of records.

(a) Suppose we have picked n entries $\{x_1,\ldots,x_n\}$. The probability that x_n is the largest is actually the same as the probability that any other entry x_i is the largest. Therefore, $\boxed{P_n = 1/n}$

(b) Since S_N is the number of records after N attempts, we may write S_N as the sum of the indicators R_i. As a result,

$$\langle S_N\rangle = \sum_{m=1}^{N}\langle R_m\rangle = \sum_{m=1}^{N} P_m = \sum_{m=1}^{N}\frac{1}{m}.$$

We may estimate the growth of $\langle S_N\rangle$ by bounding it by two integrals. Choose the integral

$$I = \int_0^N \frac{1}{n}\,dn = \ln N.$$

to be an estimate for $\langle S_N\rangle$. We see that I is an under-estimation. However, we can find how much $\langle S_N\rangle$ differs from $\ln N$ by taking a limit:

$$\lim_{N\to\infty}\langle S_N\rangle - \ln N = \texttt{EulerGamma} \approx 0.5775216\ldots$$

where `EulerGamma` is the output that I got from evaluating this limit in Mathematica. Therefore, we conclude that $\langle S_N\rangle$ grows like $\boxed{\ln N}$ in the $N \gg 1$ limit. If the number of trials double every year, then $N = N(t) = 2^t$. In this case,

$$\langle S_N\rangle \sim \ln N = \ln 2^t = \boxed{t\ln 2}$$

asymptotically.

(c) By definition, $\langle R_n R_m\rangle = \langle R_n\rangle_c\langle R_m\rangle_c + \langle R_n R_m\rangle_c = \langle R_n\rangle\langle R_m\rangle + \langle R_n R_m\rangle_c$

$$\langle R_n R_m\rangle_c = \langle R_n R_m\rangle - \langle R_n\rangle\langle R_m\rangle = P_n P_m - P_n P_m = 0.$$

(d) **(Optional)**

(e) **(Optional)**

6. Jarzynski equality.

(a) The new probability density function $p_f(W(\mu))$ is obtained from $p(\mu)$ via a change of variable transformation for which the rule is

$$p_f(W(\mu))\left|\frac{dW}{d\mu}\right| = p(\mu)$$

where $|dW/d\mu|$ is the Jacobian. Similarly, we have

$$p_b(-W(\mu'))\left|\frac{-dW}{d\mu'}\right| = p'(\mu').$$

Since there is no real thermodynamic reason for $d\mu \neq d\mu'$, we must have

$$\frac{p_f(W)}{p_b(-W)} = \frac{p(\mu)}{p'(\mu')} = \frac{\mathcal{Z}'}{\mathcal{Z}}\frac{\exp\bigl(-\beta\mathcal{H}(\mu)\bigr)}{\exp\bigl(-\beta\mathcal{H}(\mu)\bigr)\exp\bigl(-\beta W(\mu)\bigr)} = \exp\bigl[\beta(W + F - F')\bigr],$$

as desired, where we have used $\ln\mathcal{Z} = -\beta F$.

(b) Let $\Delta F = F' - F$. From Part (a) we have

$$p_f(W)e^{-\beta W} = p_b(-W)e^{-\beta\Delta F} \implies \langle e^{-\beta W}\rangle = \int p_f(W)e^{-\beta W}\,dW = e^{-\beta\Delta F}\int p_b(-W)\,dW = e^{-\beta\Delta F}.$$

As a result,

$$\Delta F = -\frac{1}{\beta}\ln\langle e^{-\beta W}\rangle = -k_BT\ln\langle e^{-\beta W}\rangle = -k_BT\ln\left[\int p_f(W)e^{-\beta W}\,dW\right],$$

as desired.

(c) We use Jensen's inequality and the fact that the function $f(W) = e^{-\beta W}$ is convex:

$$\Delta F = -k_BT\ln\langle e^{-\beta W}\rangle \leq -k_BT\ln e^{-\beta\langle W\rangle} = -k_BT\ln e^{-\langle W\rangle/k_BT} = \langle W\rangle.$$

(d) To do this problem we have to define a probability density function $\rho(\omega)$ associated with violating the second law. Given a particular W, the probability density for second law violation is $p_f(W-\omega)p_b(-W)$. Thus, $\rho(\omega)$ is in some sense a "quasi-convolution" of p_f and p_b:

$$\rho(\omega) = \int p_f(W-\omega)p_b(-W)\,dW.$$

We wish to find $\Pr(\omega > 0)$, which is given by

$$\Pr(\omega > 0) = \int_0^\infty \rho(\omega)\,d\omega.$$

In order for a factor of $e^{-\beta\omega}$ to pop up, we must make use of the fact that

$$\frac{p_f(W)}{p_b(-W)} = e^{\beta(W+F-F')}$$

In particular,

$$\left(\frac{p_f(W)}{p_b(-W)}\right)^{-1}\frac{p_f(W-\omega)}{p_b(-W+\omega)} = e^{-\beta(W+F-F')}e^{\beta(W-\omega+F-F')} = e^{-\beta\omega}.$$

As a result,

$$p_b(-W)p_f(W-\omega) = e^{-\beta\omega}p_f(W)p_b(-W+\omega) \implies \rho(\omega) = e^{-\beta\omega}\int p_f(W)p_b(-W+\omega)\,dW.$$

By shifting the integration $dW \to d(W+\omega)$, we find that

$$\rho(\omega) = e^{-\beta\omega}\int p_f(W+\omega)p_b(-W)\,d(W+\omega) = e^{-\beta\omega}\int p_f(W+\omega)p_b(-W)\,dW = e^{-\beta W}\rho(-\omega).$$

Therefore,

$$\Pr(\omega > 0) = e^{-\beta W}\int_0^\infty \rho(-\omega)\,d\omega \leq e^{-\beta W}\underbrace{\int_{-\infty}^\infty p(-\omega)\,d\omega}_{=1} \leq e^{-\beta W} = e^{-\beta W},$$

as desired, where we have used the fact that the cumulative probability is bounded above by 1.

7. (Optional) Dice.

(a) The unbiased probabilities are such that the entropy is maximized. Since 6 appears twice as many times as 1, and that entropy must be maximized, we have

$$p_6 = 2p_1 = 2a, \qquad p_2 = p_3 = p_4 = p_5 = (1-3a)/4.$$

Now we compute:

$$S = -\sum_{i=1}^{6} p_i \ln p_i = -a\ln a - 2a\ln 2a - 4\frac{1-3a}{4}\ln\frac{1-3a}{4} = -a\ln a - 2a\ln 2a - (1-3a)\ln\frac{1-3a}{4}$$

To extremize S, we find dS/da:

$$\frac{dS}{da} = -8\ln 2 + 3\ln(1-3a) - 3\log a = 3\ln\left[\frac{1-3a}{2^{8/3}a}\right] = 0 \iff 1-3a = 2^{8/3}a \implies a = p_1 = \frac{1}{2^{8/3}+3}$$

this is an local maximum because dS/da is monotonically decreasing and crosses zero at $a = 1/(2^{8/3}+3)$. We can now calculate the rest of the probabilities:

$$\boxed{p_6 = 2p_1 = \frac{2}{2^{8/3}+3}, \qquad p_2 = p_3 = p_4 = p_5 = \frac{1}{4}\left[1 - \frac{3}{2^{8/3}+3}\right] = \frac{2^{2/3}}{2^{8/3}+3}}$$

(b) The information content is the difference in entropy of a fair dice and this one.

$$I = S_{\text{fair}} - S_{\text{loaded}} = -6\times\frac{1}{6}\log_2\frac{1}{6} - \left[-a_0\log_2 a_0 - 2a_0\log_2 2a_0 - (1-3a_0)\log_2\frac{1-3a_0}{4}\right]$$

where we have defined I so that it is positive (the entropy associated with a fair dice is maximal) and $a_0 = 1/(2^{8/3}+3)$. With the help of Mathematica, we find

$$\boxed{I \approx 0.0267\ldots}$$

Mathematica code:

```
In[13]:= a0 = 1/(2^(8/3) + 3);

In[17]:= N[-Log2[1/6] - (-a0*Log2[a0] -
2*a0*Log2[2*a0] - (1 - 3*a0)*Log2[(1 - 3*a0)/4])]

Out[17]= 0.0267239
```

8. (Optional) Approach to Equilibrium

(a) By time-translation invariance we have

$$C_{ij}(t) = C_{ij}(t+\tau) = \langle x_i(t+\tau)x_j(\tau)\rangle.$$

Picking $\tau = -t$, we find

$$C_{ij}(t) = \langle x_i(0)x_j(-t)\rangle.$$

By time-reversal invariance we have

$$C_{ij}(t) = C_{ij}(-t) = \langle x_i(0)x_j(t)\rangle = \langle x_j(t)x_i(0)\rangle = C_{ji}(t),$$

as desired.

(b) As appeared in the form

$$\sqrt{\frac{\det(K)}{(2\pi)^n}}\exp\left[-\frac{1}{2}K_{ij}x_ix_j\right],$$

the matrix $[K]$ is the inverse of the covariance matrix associated with this Gaussian, i.e.,

$$[K^{-1}]_{ij} = \langle (x_i - \langle x_i\rangle)(x_j - \langle x_j\rangle)\rangle$$

On the other hand, $C_{ij}(0) = \langle x_i(0)x_j(0)\rangle$ forms the autocorrelation matrix. The two matrices are related by the well-known identity:

$$\boxed{[K]^{-1} = [C(0)] - \langle \mathbf{x}\rangle\langle \mathbf{x}\rangle^\top \implies C[0] = [K]^{-1} + \langle \mathbf{x}\rangle\langle \mathbf{x}\rangle^\top}$$

where $\langle \mathbf{x}\rangle^\top = (\langle x_1\rangle, \langle x_2\rangle \ldots)^\top$.

(c) Given $J_\alpha = -\partial \ln p(\mathbf{x})/\partial x_\alpha$, we may compute:

$$J_\alpha = -\frac{\partial}{\partial x_\alpha}\ln\left\{\sqrt{\frac{\det(K)}{(2\pi)^n}}\exp\left[-\frac{1}{2}K_{ij}x_ix_j\right]\right\} = -\frac{\partial}{\partial x_\alpha}\left[-\frac{1}{2}K_{ij}x_ix_j\right] = K_{\alpha j}x_j.$$

So,

$$\langle J_\alpha x_\beta\rangle = K_{\alpha j}\langle x_j x_\beta\rangle.$$

(d)

9. (Optional) Simpson's Paradox