1. **The climb up.**

“[The Boltzmann distribution] is the summit of statistical mechanics, and the entire subject is either the slide-down from this summit, as the principle is applied to various cases, or the climb-up to where the fundamental law is derived.” —Richard Feynman

In class we presented a reasonable but less than rigorous derivation of the Boltzmann distribution,

$$P(\nu) \propto e^{-\beta E(\nu)},$$

for microstates ν of a system that can exchange energy with a very large bath. This result was obtained from the fundamental relationship

$$P(\nu) \propto \Omega_{\rm B}(E_{\rm T} - E(\nu))$$

through a Taylor expansion of $\ln \Omega_{\rm B}$, where $E_{\rm T}$ denotes the total energy shared by system and bath, and $\Omega_{\rm B}(E_{\rm B})$ is the number of bath microstates with energy $E_{\rm B}$. Here you will develop this argument a little more carefully (and perhaps a little more convincingly—you may have wondered, for example, why we were Taylor expanding $\ln \Omega_{\rm B}$ and not $\Omega_{\rm B}$).

We will assume that $\Omega_{\rm B}$ has a large deviation form,

$$\Omega_{\rm B} = [\omega_{\rm B}(\epsilon_{\rm B})]^{N_{\rm B}},$$

where $N_{\rm B}$ is the number of molecules in the bath, $\epsilon_{\rm B} \equiv E_{\rm B}/N_{\rm B}$ is the corresponding energy per molecule, and $\omega_{\rm B}(\epsilon_{\rm B})$ is a smooth function that does not depend on the size of the bath. As in lecture the dependence of $\Omega_{\rm B}$ on $E_{\rm B}$ will be used to define a property β of the bath:

$$\beta \equiv \left(\frac{\partial \ln \Omega_{\rm B}}{\partial E_{\rm B}}\right)_{N_{\rm B},V_{\rm B}}$$

(i) Show that β is insensitive to the extent of the bath. In particular, relate β to $\omega_{\rm B}$ and derivatives of $\omega_{\rm B}$ with respect to $\epsilon_{\rm B}$. Explain why this relationship indicates independence of the bath's size. (As we will see in future lectures, this β is an inverse temperature, so it should be pleasing β doesn't depend on bath size. When I say to put a test tube in a 25° C water bath, I shouldn't have to specify if the bath is 1 liter or 2 liters in volume. Assuming the volume of the bath is very big compared to the system, I should be able to specify β alone.)

$$\beta = \left(\frac{\partial \ln \Omega_{\rm B}}{\partial E_{\rm B}}\right)_{N_{\rm B},V_{\rm B}} = N_B \frac{\partial}{\partial E_{\rm B}} \ln \omega_{\rm B}(\epsilon_{\rm B}) = \frac{N_{\rm B}}{\omega_{\rm B}}\frac{\partial \omega_{\rm B}}{\partial \epsilon_{\rm B}}\frac{\partial \epsilon_{\rm B}}{\partial E_{\rm B}} = \frac{N_{\rm B}}{\omega_{\rm B}}\frac{\partial \omega_{\rm B}}{\partial \epsilon_{\rm B}}\frac{1}{N_{\rm B}} = \frac{1}{\omega_{\rm B}}\frac{\partial \omega_{\rm B}}{\partial \epsilon_{\rm B}}$$

Notice that all factors of $N_{\rm B}$ have canceled. Since $\omega_{\rm B}$ is a smooth function that does not depend on the size of the bath, it cannot depend on $N_{\rm B}$. Increasing the system size thus yields the exact same value of β.

(ii) Show that

$$\left(\frac{\partial \Omega_{\rm B}}{\partial E_{\rm B}}\right)_{N_{\rm B},V_{\rm B}} = \beta \Omega_{\rm B}$$

$$\left(\frac{\partial \Omega_{\mathrm{B}}}{\partial E_{\mathrm{B}}}\right)_{N_{\mathrm{B}},V_{\mathrm{B}}} = \left(\frac{\partial e^{N_{\mathrm{B}} \ln \omega_{\mathrm{B}}(\epsilon_{\mathrm{B}})}}{\partial E_{\mathrm{B}}}\right)_{N_{\mathrm{B}},V_{\mathrm{B}}} = e^{N_{\mathrm{B}} \ln \omega_{\mathrm{B}}(\epsilon_{\mathrm{B}})} \frac{\partial \ln \omega_{\mathrm{B}}(\epsilon_{\mathrm{B}})}{\partial \epsilon_{\mathrm{B}}} \frac{\partial \epsilon_{\mathrm{B}}}{\partial E_{\mathrm{B}}} N_{\mathrm{B}} = \beta \Omega_{\mathrm{B}}.$$

Alternatively, you could start from the right hand side and work backwards:

$$\beta \Omega_{\mathrm{B}} = \Omega_{\mathrm{B}} \left(\frac{\partial \ln \Omega_{\mathrm{B}}}{\partial E_{\mathrm{B}}}\right) = \Omega_{\mathrm{B}} \frac{1}{\Omega_{\mathrm{B}}} \left(\frac{\partial \Omega_{\mathrm{B}}}{\partial E_{\mathrm{B}}}\right) = \frac{\partial \Omega_{\mathrm{B}}}{\partial E_{\mathrm{B}}}$$

(iii) Show that

$$\left(\frac{\partial^2 \Omega_{\mathrm{B}}}{\partial E_{\mathrm{B}}^2}\right)_{N_{\mathrm{B}},V_{\mathrm{B}}} = \beta^2 \Omega_{\mathrm{B}} + c.$$

Identify the quantity c and explain why it can be neglected in the limit $N_{\mathrm{B}} \to \infty$.

$$\left(\frac{\partial^2 \Omega_{\mathrm{B}}}{\partial E_{\mathrm{B}}^2}\right)_{N_{\mathrm{B}},V_{\mathrm{B}}} = \frac{\partial}{\partial E_{\mathrm{B}}} \left(\frac{\partial \Omega_{\mathrm{B}}}{\partial E_{\mathrm{B}}}\right)_{N_{\mathrm{B}},V_{\mathrm{B}}} = \left(\frac{\partial (\beta \Omega_{\mathrm{B}})}{\partial E_{\mathrm{B}}}\right)_{N_{\mathrm{B}},V_{\mathrm{B}}}$$
$$= \beta \left(\frac{\partial \Omega_{\mathrm{B}}}{\partial E_{\mathrm{B}}}\right)_{N_{\mathrm{B}},V_{\mathrm{B}}} + \Omega_{\mathrm{B}} \left(\frac{\partial \beta}{\partial E_{\mathrm{B}}}\right)_{N_{\mathrm{B}},V_{\mathrm{B}}} = \beta^2 \Omega_{\mathrm{B}} + c,$$

where

$$c = \Omega_{\mathrm{B}} \left(\frac{\partial^2 \ln \Omega_{\mathrm{B}}}{\partial E_{\mathrm{B}}^2}\right)_{N_{\mathrm{B}},V_{\mathrm{B}}} = \Omega_{\mathrm{B}} \left(\frac{\partial \beta}{\partial E_{\mathrm{B}}}\right)_{N_{\mathrm{B}},V_{\mathrm{B}}} = \Omega_{\mathrm{B}} \frac{\partial \beta}{\partial \epsilon_{\mathrm{B}}} \frac{\partial \epsilon_{\mathrm{B}}}{\partial E_{\mathrm{B}}} = \frac{\Omega_{\mathrm{B}}}{N_{\mathrm{B}}} \frac{\partial \beta}{\partial \epsilon_{\mathrm{B}}}.$$

Because neither β nor ϵ_{B} depend on the system size N_{B}, the derivative on the right hand side also does not depend on N_{B}. The ratio of $c/(\beta^2 \Omega_B)$ decays as $1/N_{\mathrm{B}}$, and hence c can be neglected in the $N_{\mathrm{B}} \to \infty$ limit.

(iv) Calculate

$$\left(\frac{\partial^n \Omega_{\mathrm{B}}}{\partial E_{\mathrm{B}}^n}\right)_{N_{\mathrm{B}},V_{\mathrm{B}}}$$

for arbitrary (integer) n. Do not include any terms that are negligible in the limit $N_{\mathrm{B}} \to \infty$.

You can continue using the chain rule as in (iii) and observe that all terms involving a derivative of β with respect to E_{B} will have contributions which are negligible in the $N_{\mathrm{B}} \to \infty$ limit. That leaves only

$$\lim_{N_{\mathrm{B}} \to \infty} \left(\frac{\partial^n \Omega_{\mathrm{B}}}{\partial E_{\mathrm{B}}^n}\right)_{N_{\mathrm{B}},V_{\mathrm{B}}} = \beta^n \Omega_{\mathrm{B}}.$$

(v) Consider the Taylor expansion

$$\begin{aligned}\Omega_{\rm B}(E_{\rm T}-E) &= \Omega_{\rm B}(E_{\rm T}) - E\left(\frac{\partial \Omega_{\rm B}}{\partial E_{\rm B}}\right)_{N_{\rm B},V_{\rm B}} + \frac{1}{2}E^2\left(\frac{\partial^2 \Omega_{\rm B}}{\partial E_{\rm B}^2}\right)_{N_{\rm B},V_{\rm B}} + \dots \\ &= \sum_{n=0}^{\infty}\frac{1}{n!}(-E)^n\left(\frac{\partial^n \Omega_{\rm B}}{\partial E_{\rm B}^n}\right)_{N_{\rm B},V_{\rm B}},\end{aligned}$$

where all partial derivatives are implicitly evaluated at $E = 0$.

Combining your results together with this expansion, show that the relationship

$$\Omega_{\rm B}(E_{\rm T} - E(\nu)) \propto e^{-\beta E(\nu)}$$

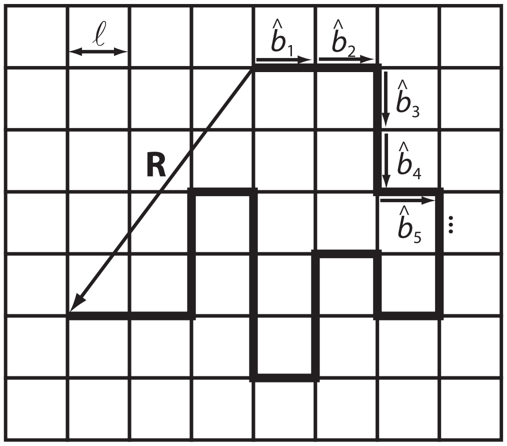
is exact in the limit $N_{\rm B} \to \infty$ of an infinitely large bath.

$$\begin{aligned}\Omega_{\rm B}(E_{\rm T}-E) &= \sum_{n=0}^{\infty}\frac{1}{n!}(-E)^n\left(\frac{\partial^n \Omega_{\rm B}}{\partial E_{\rm B}^n}\right)_{N_{\rm B},V_{\rm B}} \\ &= \sum_{n=0}^{\infty}\frac{1}{n!}(-\beta E)^n\Omega_{\rm B}(E_{\rm T}) \\ &= \Omega_{\rm B}(E_{\rm T})e^{-\beta E}\end{aligned}$$

For the step connecting the last two lines above, we pull $\Omega_{\rm B}(E_T)$ out in front of the sum (since it is not dependent on n). We also remember that the Taylor Expansion around $x = 0$ for e^{ax} (where a is a constant) is $e^{ax} = \sum_{n=0}^{\infty}\frac{1}{n!}a^n e^{0a} = \sum_{n=0}^{\infty}\frac{1}{n!}a^n$. This appears above where $a = (-\beta E)$.

2. **A simple model of a polymer.**

You're no doubt aware that the shape a protein folds up into is influenced by energetic interactions between amino acids. You may not have given as much thought to the influence of entropy on a protein's shape. To focus on that concept, we consider the simplest model for conformational fluctuations of a long chain molecule, formed from n polymer segments connected end to end. Each segment $i = 1, 2, \dots, n$ (perhaps comprising many chemical units) has a fixed length ℓ and an orientation $\hat{b}_i$ that is parallel to one of d Cartesian axes ($\hat{x}$, $\hat{y}$, or $\hat{z}$ in 3 dimensions). In other words, the molecular configuration traces a random walk on a d-dimensional cubic lattice:



Imagine that the orientations of different segments are statistically independent, and that there is no preferred orientation, $\langle \hat{b}_i \rangle = 0$ and $\langle \hat{b}_i \cdot \hat{b}_j \rangle = \delta_{ij}$, where $\delta_{ij} = 1$ if $i = j$ and vanishes otherwise.

(i) Show that the number of microstates of such an ideal chain molecule has the form $\Omega = a^n$. Determine the parameter a as a function of dimensionality d. (Assume that every chain starts at the origin.)

The molecule consists of n bonds, each of which can be selected from $2d$ options. Hence the total number of possible polymers is $\Omega = (2d)^n$. The entropy (excluding any translational entropy from the location of the first monomer that I'm fixing at the origin) is therefore $S = k_\mathrm{B} \ln \Omega = k_\mathrm{B} n \ln 2d$.

(ii) What is the expected value of $\boldsymbol{R}$, that is to say what is $\langle \boldsymbol{R} \rangle$, where $\boldsymbol{R} = \ell \sum_{i=1}^n \hat{b}_i$ is the end-to-end vector?

By symmetry, $\langle \boldsymbol{R} \rangle = 0$. Every step is as likely to go left as to go right, up as to go down, etc. Looking ath the expression for $\boldsymbol{R}$, we can also show that $\langle \boldsymbol{R} \rangle = \left\langle \ell \sum_{i=1}^n \hat{b}_i \right\rangle = \ell \sum_{i=1}^n \left\langle \hat{b}_i \right\rangle = \ell \sum_{i=1}^n 0$.

(iii) The squared length of the end-to-end vector is $R^2 = \boldsymbol{R} \cdot \boldsymbol{R}$. Calculate the mean squared end-to-end distance of the chain molecule, $\left\langle R^2 \right\rangle$. Your result should indicate that the typical distance between ends of the molecule grows with chain length as $\sqrt{\langle R^2 \rangle} \propto n^\nu$. Identify the exponent ν. How does your result depend on the dimensionality d?

Notice that $\boldsymbol{R} = \ell \sum_{i=1}^n \hat{b}_i$, so

$$\begin{aligned}
\langle \boldsymbol{R} \cdot \boldsymbol{R} \rangle &= \ell^2 \langle \sum_{i=1}^n \sum_{j=1}^n \hat{b}_i \cdot \hat{b}_j \rangle \\
&= \ell^2 \sum_{i=1}^n \left(\langle \hat{b}_i \cdot \hat{b}_i \rangle + \sum_{j \neq i} \langle \hat{b}_i \cdot \hat{b}_j \rangle \right) \\
&= \ell^2 \sum_{i=1}^n \langle \hat{b}_i \cdot \hat{b}_i \rangle \\
&= \ell^2 \sum_{i=1}^n 1 \\
&= \ell^2 n.
\end{aligned}$$

Thus, $\sqrt{\langle R^2 \rangle} = \ell n^{1/2}$. Remarkably the answer does not depend on the dimensionality d!

(iv) For a (three-dimensional) polymer in "good" solvent, experiments yield $\nu \approx 3/5$. (A "good" solvent is one that prevents the molecule from collapsing onto itself, i.e., effective interactions among different segments are repulsive.) Compare this measured value with the one you calculated, and comment on the discrepancy.

The three-dimensional polymer in good solvent is slightly more spread out than what we got from our model. The difference seems to be that our polymer can pass back over itself. If, for example, we changed the rules so the polymer could not overlap itself (as the good solvent apparently does), then we should anticipate less backtracking and more spread.

3. **Exploring the lattice polymer with a computer**

One way to get more insight into the previous problem is to *sample* realizations of the polymer using a computer. To illustrate the idea, let's temporarily think about last week's coin flip problem. With 3 coin flips, one can enumerate (that is list out) all of the possible sequences of coin flips: HHH, HHT, HTH, THH, HTT, THT, TTH, TTT. You could even write a computer program to systematically generate the possibilities, but for N coin flips there will be 2^N possibilities. As N gets bigger and bigger, it becomes less and less reasonable to use a computer to generate all of the possibilities. Rather, it is commonly useful to generate *representative samples* from the probability distribution. You might, for example, use a computer to randomly flip 100 coins in a row then use that particular 100-flip sequence as a good guess for what you would expect. Since all of the coin flip microstates are equally likely, it is kind of crazy to think that any one of your randomly generated sequences is actually a good guide for what to expect. Nevertheless, you could collect several different 100-flip sequences and take that collection of samples as a good reflection of the distribution of possibilities. In particular, you could approximate the distribution for a macrostate, like the total number of times you got heads, by constructing a histogram that counts up how many times your samples returned 0 heads, 1 head, 2 heads, etc.

For the polymer problem, we generate representative samples of these polymers by randomly picking a direction in which to extend the polymer during each step, with all of the directions on the grid being equally probable. Online, you will find a link to some code that generates such random polymers in both two- and three-dimensional space.

(i) Using the provided code (or by writing your own), generate some samples of the random lattice polymer in both two and three dimensions with $n = 1000$. Is it more common to see the polymer with bunched up regions that appear as though it is partially "folded" or more likely to see the polymer completely spread out as if it is "unfolded"? Notice that there is nothing like a Coulomb attraction pulling any of the monomers closer together. What you are seeing is entirely entropic (due to counting the number of possibilities).

The sample has some sections that look bunched up and "folded". For the polymer to really be fully spread out requires most of the steps to continue in the same direction, but it's far more likely that steps will double back on themselves. This doubling back just occurs because there are more possibilities to do so, not because there is something like an electrostatic attraction pulling the polymer into a folded state.

(ii) Generate 100 different $n = 1000$ 3d polymers. How many of the 100 samples have the polymer start and stop at the origin? Does this observation contradict your response to 2(ii)?

I think it will be exceedingly unlikely to see the polymer start and stop at the origin. It is true that of all the end-to-end vectors, $\mathbf{0}$ is the most likely, and it's also true that the average end-to-end vector is $\mathbf{0}$, but the average end-to-end distance is not zero. The difference is that when averaging the end-to-end separation, a value of -5 cancels with a value of 5 since symmetry tells us each are equally likely. However, both such polymers have an end-to-end *distance* of 5. It turns out that it is far more likely to find a nonzero end-to-end distance than it is to find 0.

(iii) Given your response to part (ii), suggest a reason that the root mean squared end-to-end distance $\sqrt{\langle \boldsymbol{R}^2 \rangle}$ would be a better measure of the size of the polymer than the average displacement vector $\langle \boldsymbol{R} \rangle$.

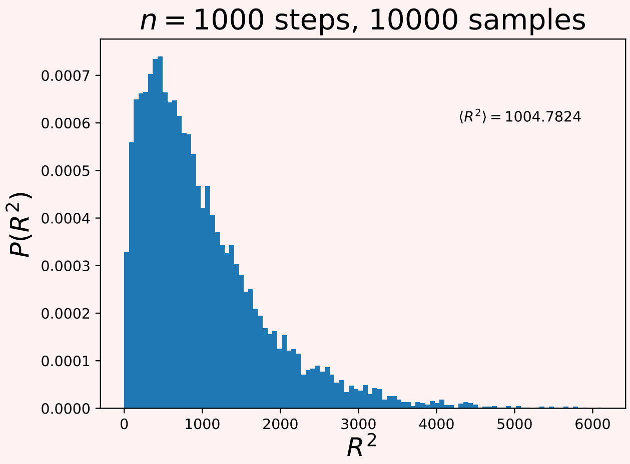
The root mean squared end-to-end distance circumvents the cancellations described in part v. It gives a typical breadth of the distribution, even when the distribution of displacements is centered at zero.

(iv) Using sufficiently many samples of the $n = 1000$ polymer in three dimensions, plot a histogram to approximate the probability distribution for R^2. Does your histogram agree with your answer to 2(iii)? Discuss.

The following python code can generate the histogram:

```
import matplotlib as mpl
import matplotlib.pyplot as plt
import numpy as np
plt.style.use(’default’)

# Each step can be either right, left, up, or down
StepChoices3D = np.array([[0, 1, 0], [0, -1, 0], [1, 0, 0], [-1, 0, 0],
                         [0, 0, 1], [0, 0, -1]])

# Generate a vector of Num_Steps random numbers,
# each drawn uniformly between 0 and 1
Num_Steps = 1000

Num_Samples = 10000
R2values = np.zeros(Num_Samples)

for s in range(Num_Samples):
  # generate the random numbers for polymer number s
  RandomNumbers = np.random.random(Num_Steps)

  # Initialize the 3d polymer at the origin
  polymer3D = np.zeros([Num_Steps + 1, 3])

  # Compute the position of the monomers from RandomNumbers
  for i in range(Num_Steps):
    polymer3D[i+1] = polymer3D[i] + \
      StepChoices3D[int(np.floor(RandomNumbers[i]*6))]

  # Record the distance between endpoints of the polymer
  R2values[s] = np.sum((polymer3D[Num_Steps] - polymer3D[0])**2)

plt.hist(R2values,100,density=True);
plt.xlabel(r’$R^2$’, fontsize=18);
plt.ylabel(r’$P(R^2)$’, fontsize=18);
plt.title(r’$n = $’+str(Num_Steps)+’ steps, ’+
          str(Num_Samples)+’ samples’, fontsize=20);
plt.text(np.max(R2values)*0.7, 0.0006, r’$\langle R^2 \rangle =’ +
```

```
         str(np.mean(R2values))+'$');
plt.tight_layout()
plt.savefig('r2dist.eps',format='eps',transparent=True);
```



Notice that the average is (to within sampling error) equal to $Nl^2 = 1000$, which is what we expect.

4. **Fun with Gaussians.** On the last problem we had several occasions to confront the Gaussian distribution,

$$\rho(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

Some students expressed that it would be useful to build a little more familiarity with the distribution. Consider these exercises to be either a crash course or a refresher. I do not consider it important that you "discover" the standard tricks for computing the integrals, but I do think you should be aware of them. Please talk to me or to other students if you're having trouble figuring out how to proceed!

(i) **Normalization:** Show that $\int dx\, \rho(x) = 1$. [*Standard trick: Convert to a two-dimensional integral over a joint distribution of identical, independent Gaussians for the x and y coordinates. Convert to polar coordinates, taking care to include the appropriate Jacobian factor.*]

Consider the distribution

$$\begin{aligned} P(x,y) &= \left[\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\right]\left[\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right)\right] \\ &= \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(x-\mu)^2+(y-\mu)^2}{2\sigma^2}\right). \end{aligned}$$

Let $\boldsymbol{z} = (x,y)$ be the vector of x and y coordinates and $\boldsymbol{\mu} = (\mu,\mu)$. We now show that

$$P(\boldsymbol{z}) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(\boldsymbol{z}-\boldsymbol{\mu})^2}{2\sigma^2}\right)$$

is a normalized distribution. We change variables, setting $\boldsymbol{u} = \boldsymbol{z} - \boldsymbol{\mu}$, with $d\boldsymbol{z} = d\boldsymbol{u}$, so

$$\begin{aligned}
\int d\boldsymbol{z}\, P(\boldsymbol{z}) &= \int d\boldsymbol{u}\, \frac{1}{2\pi\sigma^2} \exp\left(-\frac{\boldsymbol{u}^2}{2\sigma^2}\right) \\
&= \int_0^\infty dr\, r \int_0^{2\pi} d\theta\, \frac{1}{2\pi\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right), \text{ switching to polar coordinates} \\
&= \frac{1}{\sigma^2} \int_0^\infty dr\, r \exp\left(-\frac{r^2}{2\sigma^2}\right) \\
&= \frac{1}{\sigma^2} \int_0^\infty \sigma^2 dv\ \exp(-v), \text{ where } v = r^2/2\sigma^2 \\
&= -\exp(-v)|_0^\infty \\
&= 1.
\end{aligned}$$

It remains to show that $P(\boldsymbol{z})$ being normalized implies that $P(x)$ is also normalized. Since x and y are independent random variables, $P(\boldsymbol{z}) = P(x, y) = P(x)P(y)$. Integration yields

$$\begin{aligned}
1 &= \int d\boldsymbol{z}\, P(\boldsymbol{z}) \\
&= \left(\int_{-\infty}^{\infty} dx\, P(x)\right) \left(\int_{-\infty}^{\infty} dy\, P(y)\right) \\
&= \left(\int_{-\infty}^{\infty} dx\, P(x)\right)^2,
\end{aligned}$$

which does the trick.

(ii) **Mean:** Show that $\langle x \rangle = \int_{-\infty}^{\infty} dx\, xP(x) = \mu$. [*Standard trick: Substitute $u = x - \mu$ and notice that an integral cancels by symmetry.*]

$$\begin{aligned}
\langle x \rangle &= \int_{-\infty}^{\infty} dx\, \frac{x}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \\
&= \int_{-\infty}^{\infty} du \frac{u+\mu}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{u^2}{2\sigma^2}\right), \text{ where } u = x - \mu \\
&= \mu + \int_{-\infty}^{\infty} du\, \frac{u}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{u^2}{2\sigma^2}\right) \\
&= \mu.
\end{aligned}$$

The final integral cancels by symmetry.

(iii) **Variance:** Show that $\left\langle (\delta x)^2 \right\rangle = \int_{-\infty}^{\infty} dx\, (x-\mu)^2 P(x) = \sigma^2$. [*Standard trick: From (i) and symmetry find the integral of $e^{-\alpha x^2}$ from 0 to ∞. Differentiate with respect to α.*]

Notice that

$$\int_0^\infty dx\; e^{-\alpha x^2} = \frac{1}{2}\sqrt{\frac{\pi}{\alpha}},$$

using (i) and symmetry. You could also reason this from the standard normalization of a Gaussian, which guarantees that $\sqrt{2\pi\sigma^2} = \int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2\sigma^2}\right)$ and where $\alpha = \frac{1}{2\sigma^2}$. Now we differentiate both sides of the equality with respect to α. (Notice that in this step we also multiplied both sides by -1):

$$-\frac{\partial}{\partial\alpha}\left(\int_0^\infty dx\; e^{-\alpha x^2}\right) = \int_0^\infty dx\; x^2 e^{-\alpha x^2} = -\frac{\partial}{\partial\alpha}\left(\frac{1}{2}\sqrt{\frac{\pi}{\alpha}}\right) = \frac{1}{4}\sqrt{\frac{\pi}{\alpha^3}}.$$

The integral from 0 to ∞ gives the same result as the integral from $-\infty$ to 0, so we have

$$\sqrt{\frac{\alpha}{\pi}}\int_{-\infty}^{\infty} dx\; x^2 e^{-\alpha x^2} = \frac{1}{2\alpha}.$$

Substituting in $2\sigma^2 = 1/\alpha$ gives

$$\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} dx\; x^2 e^{-\alpha x^2} = \sigma^2.$$

The final step is to confirm that a shift $u = x - \mu$ does not change the value of the integral. Notice that $du = dx$ and that the infinite bounds are unchanged by a shift by μ.