Maxwell-Boltzmann Distribution   *by Tony Shing*


Overview:

- Prerequisite: Probability, extending to continuous random variables

- Thermal fluctuation and Boltzmann factor

- Maxwell-Boltzmann distribution, average speed of ideal gas

---

1 Probability Distribution: Discrete & Continuous

1.1 Probabilty Mass Function (PMF)

For random variables that can be picken from a finite number cases, we describe them using the **probability mass function** (PMF), $P(x)$, or most of the time we just call them probabilities.

E.g. By throwing 2 dices and sum their values, we obtain a list of different probabilities for getting any integer between 2 to 12.
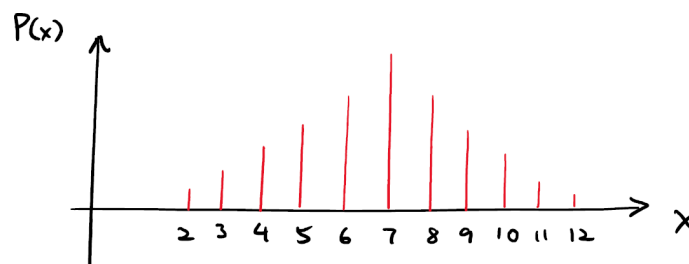
- **Normalization requriement**: The probabilities of all possible outcomes add up to 1.

$$\sum_{\text{all possible } x_i} P(x_i) = 1$$

- **Expected value**: i.e. The weighted average of all outcomes according to the probabilities, which is the average value you would get after repeating the process ∞ times.

$$\mathrm{E}[x] \stackrel{\text{def}}{=} \sum_{\text{all possible } x_i} x_i P(x_i)$$

- **Variance**: It is the average of $(\text{distance})^2$ of all outcomes relative to the expected value. Taking square root gives the **standard derivation** (S.D.) $= \sqrt{\text{Variance}}$.

$$\mathrm{Var}[x] \stackrel{\text{def}}{=} \underline{\mathrm{E}[(x - \mathrm{E}[x])^2]}$$

Then take average $\longrightarrow$ $\qquad$ $(x - \mathrm{E}[x])$ = x's distance from the expected value
Square for taking magnitude only

1.2 Probability Density Function (PDF)

For random variables that can apprear as any value within an interval, we describe them using the **probability density function** (PDF).
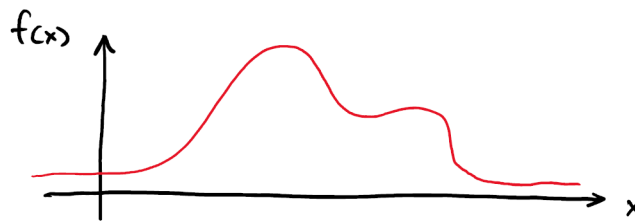
Pay attention here: the output value of the PDF, $f(x)$, is NOT the probability of getting x. This is because there are always infinitely many real number along any interval of real numbers. The probability of getting an exact number is limited to 0.

Instead, PDF can only describe the probability of getting a value within an interval $[x_0, x_0 + \mathrm{d}x]$.

$$P(\underbrace{x_0 < x < x_0 + \mathrm{d}x}_{\text{Prob. inside the interval}}) = \underbrace{f(x)}_{\text{height}}\underbrace{\mathrm{d}x}_{\text{width}}$$
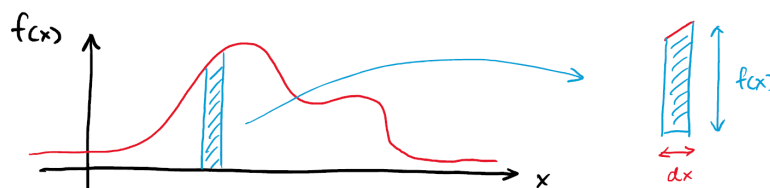
As interpreted, the probability is represented by the area under curve $f(x)\,\mathrm{d}x$, not the height $f(x)$. This is also why it is named "density" function - the probability per "length" of x.

The formula for PMF (discrete cases) can be extended to PDF (continuous cases).

- **Normalization requrement**: The probabilities of all possible cases add up to 1. From the condition for PMF, we can extend the sum to integral for PDF.

$$\sum_{\text{all possible } x_i} P(x_i) = 1 \qquad \Rightarrow \qquad \int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1$$

- **Expected value**: The weighted average of all outcomes according to the probabilities. Can be understand by the weighted sum interpretation of integration.

$$\mathrm{E}[x] \overset{\text{def}}{=} \sum_{\text{all possible } x_i} x_i P(x_i) \qquad \Rightarrow \qquad \mathrm{E}[x] \overset{\text{def}}{=} \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x$$

- **Variance**: Same as in the discrete case, it is the average of (distance)^2 of all outcomes relative to the expected value.

$$\mathrm{Var}[x] \overset{\text{def}}{=} \mathrm{E}\big[(x - \mathrm{E}[x])^2\big] \qquad \Rightarrow \qquad \mathrm{Var}[x] \overset{\text{def}}{=} \int_{-\infty}^{\infty} (x - \mathrm{E}[x])^2 f(x)\,\mathrm{d}x$$

Taking square root gives the standard derivation (S.D.) $= \sqrt{\mathrm{Var}[x]}$.

---

2 Fluctuation & Equilibrium

2.1 Fluctuation of State Parameters

From previous notes, we have learnt that

- Different states of a thermodynamics system are distinguished by macrostate parameters.

- Change of macrostate is caused by random motion.

- Every *closed* system will eventually evolve into a state with maximum multiplicity, because the probability to observe a macrostate is proportional to its multiplicity.

In general, if we can express the entropy S of system in terms of a macrostate parameter x, We can compute the corresponding $x = x_0$ at the maximum entropy state by solving

$$\begin{cases} \dfrac{\mathrm{d}S}{\mathrm{d}x}\bigg|_{x=x_0} = 0 & \left(\substack{\text{Entropy is at its maximum} \\ \Rightarrow \text{ Slope} = 0 \text{ at } x=x_0}\right) \\[2ex] \dfrac{\mathrm{d}^2 S}{\mathrm{d}x^2}\bigg|_{x=x_0} < 0 & \left(\substack{\text{Slope keeps decreasing} \\ \text{near the maximum}}\right) \end{cases}$$

However, because of random motion, even if the system has reached its maximum entropy state, its state parameters can still fluctuate. Every so often, the parameter x may deviate by some small amount Δx, causing the system entering a lower entropy state. But this change will revert very soon again due to random motion.

For example, in the 2-boxes system with 1000 balls, its macrostates can be distinguished by one state parameter - N_R = the number of balls in the right box. Because the balls are free to move (randomly) between the boxes, N_R must fluctuate, even when it has reached the maximum entropy state 500L-500R.

- One of the ball may temporarily move from the left box to the right box. The system's state change to 499L-501R.

- After some time, another ball moves from the right box to the left box. The system's state returns to 500L-500R.

<center>(add figure here: moving ball in and out)</center>

In fact, we can determine the probability distribution of observing fluctuation Δx. Recall that the probability of observing a macrostate with parameter x is proportional to the number of microstates (i.e. multiplicity) it corresponds to,

$$P(x = x_0) \ \propto \ \left(\begin{array}{c} \text{\# of microstates that} \\ \text{leads to observing } x = x_0 \end{array} \right)$$

$$= \frac{W(x = x_0)}{Z} = \frac{e^{\frac{S(x=x_0)}{k}}}{Z}$$

where Z is just some constant for normalization (because probability should add up to 1).

$$Z = \begin{cases} \displaystyle\sum_{\substack{\text{all possible} \\ x_i}} W(x_i) = \sum_{\substack{\text{all possible} \\ x_i}} e^{\frac{S(x_i)}{k}} & \text{if } x \text{ must take discrete values} \\[2em] \displaystyle\int_{\substack{\text{possible} \\ \text{range of } x}} W(x) \, \mathrm{d}x = \int_{-\infty}^{\infty} e^{\frac{S(x)}{k}} \, \mathrm{d}x & \text{if } x \text{ can take continuous values} \end{cases}$$

Let $x = x_0$ when the system is at its maximum entropy state, the fluctuation in entropy ΔS can be related to the fluctuation of the state parameter Δx by Taylor series expansion:

$$S(x_0 + \Delta x) = S(x_0) + \Delta S$$

$$= \underbrace{S(x_0)}_{\substack{\text{A constant}}} + \underbrace{\left.\frac{\mathrm{d}S}{\mathrm{d}x}\right|_{x=x_0}}_{\substack{S = \text{max.} \\ \text{so slope} = 0}} (\Delta x) + \frac{1}{2} \underbrace{\left.\frac{\mathrm{d}^2 S}{\mathrm{d}x^2}\right|_{x=x_0}}_{\substack{S = \text{max.} \\ 2^{\text{nd}} \text{ derivative} < 0}} (\Delta x)^2 + \cdots$$

The probability distribution of $x = x_0 + \Delta x$ is therefore approximately

$$P(x) = \frac{e^{\frac{S(x)}{k}}}{Z}$$

$$\approx \underbrace{\frac{1}{Z} \cdot e^{\frac{S(x_0)}{k}}}_{\substack{\text{Both are constant} \\ \text{can combine}}} \cdot e^{\frac{1}{2k}\frac{\mathrm{d}^2 S}{\mathrm{d}x^2}\big|_{x_0}(\Delta x)^2}$$

$$= \frac{1}{Z'} \cdot e^{\frac{1}{2k}\frac{\mathrm{d}^2 S}{\mathrm{d}x^2}\big|_{x_0}(x-x_0)^2}$$

<center>P.4</center>

This probability is in the form of **normal distribution** (also called Gaussian distribution), whose standard formula is written as

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

A normal distribution is governed by two variables, μ and σ, which are the mean and standard derivation of the distribution respectively.

(add figure here: gaussian distribution)

By comparison, we can identify that the distribution of x follows a normal distribution with

- Mean: $\mu = x_0$
  As expected, x should fluctuate around $x = x_0$, which is the value of x when the system is in the maximum entropy state.

- Standard derivation: $\sigma = -\left(\frac{1}{k}\frac{\mathrm{d}^2 S}{\mathrm{d}x^2}\big|_{x_0}\right)^{-\frac{1}{2}}$
  i.e. The size of fluctuation depends on the 2nd derivative of S with respect to x.

Side note:

Recall in Clausius theorem - entropy of a closed system must either be a constant or increase along any process. But because macrostate parameters are always fluctuating due to random process, entropy can fluctuate too. Clausius theorem should be more accurately written as:

$$\Delta S \geq 0 \quad \xRightarrow{\text{More accurately}} \quad \Delta S_{\text{avg.}} \geq 0$$

2.2 Thermal Equilibrium

In a closed system, the total energy must conserve within the system. However because thermal fluctuation is allowed, the internal energy of individual objects in the system are not constants but fluctuate around some values.

(add figure here: 2 object exchange heat)

For simplicity, we can consider a system with only 2 objects A and B. Because the total energy U_{total} is conserved, if A has an internal energy U_A, B must have an internal energy $U_B = U_{\text{total}} - U_A$.

Solely using U_A is enough to distinguish different states.
We can investigate the fluctuation of U_A as a state parameter.

Suppose the two objects' entropy can be expressed in terms of their internal energy:

$$S_A = S_A(U) \quad \text{and} \quad S_B = S_B(U)$$

We can determine at what condition when the system is at the maximum entropy state. Because entropy of a system is the sum of entropy of individual objects,

$$S_{\text{total}} = S_A(U_A) + S_B(U_B) = S_A(U_A) + S_B(U_{\text{total}} - U_A)$$

At the maximum entropy state, $\dfrac{\mathrm{d}}{\mathrm{d}U_A} S_{\text{total}} = 0$.

$$\frac{\mathrm{d}}{\mathrm{d}U_A} S_{\text{total}} = \frac{\mathrm{d}S_A}{\mathrm{d}U_A} + \frac{\mathrm{d}S_B}{\mathrm{d}U_A}$$

$$= \frac{\mathrm{d}S_A}{\mathrm{d}U_A} + \frac{\mathrm{d}S_B}{\mathrm{d}U_B} \cdot \frac{\mathrm{d}}{\mathrm{d}U_A}(U_{\text{total}} - U_A)$$

$$= \frac{\mathrm{d}S_A}{\mathrm{d}U_A} - \frac{\mathrm{d}S_B}{\mathrm{d}U_B}$$

$$= 0$$

$$\frac{1}{T_A} \overset{\text{def}}{=} \frac{\mathrm{d}S_A}{\mathrm{d}U_A} = \frac{\mathrm{d}S_B}{\mathrm{d}U_B} \overset{\text{def}}{=} \frac{1}{T_B}$$

This result implies - If the objects' entropy-energy relation $S(U)$ is known, we can define a state parameter $T = \dfrac{\mathrm{d}U}{\mathrm{d}S}$ (so-called "temperature"), which can be used to indicate **thermal equilibrium**:

> Thermal equilibrium occurs at the maximum entropy state of a system, when every object's "T" parameter must be equal.

Remember: Thermal equilibrium only requires objects to have equal temperature. Fluctuation in internal energy is still allowed. Since the 2$^{\text{nd}}$ derivative of S by U_A is

$$\frac{\mathrm{d}^2 S}{\mathrm{d}U_A^2} = \frac{\mathrm{d}}{\mathrm{d}U_A}\left(\frac{1}{T}\right) = -\frac{1}{T^2}\frac{\mathrm{d}T}{\mathrm{d}U_A} = -\frac{1}{T^2 C_V}$$

where C_V is the heat capacity of object A under constant volume process. So the distribution of U_A is a normal distribution:

- Centered around some equilibrium value U_0.
  (we cannot solve the value without knowing what exactly are $S_A(U)$ and $S_B(U)$.)

- Standard derivation $= T\sqrt{kC_V}$.
  This is the characteristic size of energy fluctuation of a system.

(add figure here: normal distribution with standard derivation: 95% lies within 2SD)

How large is this energy fluctuation? Asumming room temperature $T \sim 300\text{K}$,

- In daily life scale:

  - Heat capacity of materials $C_V \sim 10^3 \text{J/K}$.
  - Energy fluctuation on objects is therefore of magnitude $\sim 10^{-10}\text{J}$

which is completely negligible in our daily life.

- In atomic scale:
    - One ideal gas particle has a heat capacity $C_V = \frac{3}{2}k$.
    - Energy fluctuation on one particle is therefore $T\sqrt{kC_V} \sim kT \sim 10^{-21}\text{J}$.
    - Mass of a proton/neutron is in the scale $m \sim 10^{-27}\text{kg}$.

A sudden fluctuation in energy is enough to accelerate a particle's speed by $\sim 1000\text{m/s}$!

$$\Delta U = kT \approx \frac{1}{2}m(\Delta v)^2 \sim 10^{-21} \qquad \Rightarrow \qquad \Delta v = \sqrt{\frac{kT}{m}} \sim 10^3$$

This tells us what actually happens in thermal equilibrium:

> The concept of thermal equilibrium is "sensible" only for objects in large scale,
> so that any effects of energy fluctuation are averaged out by random motion.
> But in microscopic scale, the particles are interacting violently all the time.

(add figure here: normal scale: quiet; microscale: crazy)

2.3 Boltzmann Factor

Because of the scale of energy fluctuation is so big for a particle, we cannot expect the internal energy of a particle to always stay around the value which maximize the entropy of the object. (It is the average U that maximizes the entropy, not the U on a single particle.)

What is the probability for a particle having a specific energy U under thermal equilibrium?

We can model by isolating a particle with the rest of every particle in a box. Although the particle may be moving crazily due to thermal fluctuation, we know that the rest of the particles are "collectively" forming a thermal equilibrium of temperature T.

(add figure here: heat bath model: bath = rest of particle, temp T; obj = 1 particle, no T defined)

Let the total energy to be a constant U_{total}. Obviously, the energy carried by 1 particle U_p should be way less than that carried by the rest U_{rest}. We can write

$$U_{\text{total}} = U_{\text{rest}} + U_p = \text{const.} \qquad \text{and} \qquad U_{\text{rest}} \gg U_p$$

Let the entropy of the rest of the particle be some function $S_{\text{rest}} = S_{\text{rest}}(U_{\text{rest}})$. We can make a Taylor expansion by

$$\begin{aligned} S_{\text{rest}} &= S_{\text{rest}}(U_{\text{rest}}) \\ &= S_{\text{rest}}(U_{\text{total}} - U_p) \\ &\approx S_{\text{rest}}(U_{\text{total}}) - \left.\frac{\mathrm{d}S_{\text{rest}}}{\mathrm{d}U}\right|_{U=U_{\text{total}}} \cdot U_p \\ &\approx S_{\text{rest}}(U_{\text{total}}) - \frac{U_p}{T} \end{aligned}$$

The total multiplicity of the box is thus

$$S_{\text{total}} = S_{\text{rest}} + S_p$$

$$W_{\text{total}} = e^{\frac{S_{\text{rest}} + S_p}{k}}$$

$$\approx e^{\frac{S_{\text{rest}}(U_{\text{total}}) - \frac{U_p}{T}}{k}} \cdot e^{\frac{S_p}{k}}$$

$$= e^{\frac{S_{\text{rest}}(U_{\text{total}})}{k}} \cdot e^{-\frac{U_p}{kT}} \cdot e^{\frac{S_p}{k}}$$

$$= \underbrace{e^{\frac{S_{\text{bath}}(U_{\text{total}})}{k}}}_{\substack{U_{\text{total}} \text{ is a constant} \\ \text{so this term is just a constant}}} \cdot e^{-\frac{U_p}{kT}} \cdot \underbrace{W_p(U_p)}_{\substack{\text{Multiplicity of the particle} \\ \text{is some function of } U_p}}$$

Now the total multiplicity is a function of U_p only. Thus we can calculate the probability that the particle acquire an internal energy U_p:

$$P(U = U_p) \ \propto \ \left(\begin{array}{c} \text{\# of microstates that} \\ \text{leads to observing } U = U_p \end{array} \right)$$

$$= W_{\text{total}}(U_p)$$

$$\propto e^{-\frac{U_p}{kT}} \cdot W_p(U_p)$$

Upon normalizing, we define the **Boltzmann distribution** - the probability of finding a particle with internal energy U within an object/system that is in thermal equilibrium of temperature T.

$$\boxed{P(U) = \frac{e^{-\frac{U}{kT}} \cdot W_p(U)}{Z}}$$

The components of the Boltzmann distribution also have some fancy names:

- The term $e^{-\frac{U}{kT}}$ is called the **Boltzmann factor**.

- The normalization constant Z is called **partition function**, which calculates as

$$Z = \begin{cases} \displaystyle\sum_{\substack{\text{all possible} \\ \text{values of } U}} e^{-\frac{U}{kT}} \cdot W_p(U) & \text{if } U \text{ must take discrete values} \\[2em] \displaystyle\int_{\substack{\text{possible} \\ \text{range of } U}} e^{-\frac{U}{kT}} \cdot W_p(U) \, \mathrm{d}U & \text{if } U \text{ can take continuous values} \end{cases}$$

Example 2.1. Suppose now the fluctuation is not on one particle, but four particles - which these four particles is contained in a 2-boxes system, and particles in the right box would have higher potential energy than in the left box by ϵ.

(add figure here: bath model -¿ connect to 2 boxes model -¿ show energy diagram)

Given that the thermal equilibrium side is at temperature T. What is the probability to find each of the macrostate of 2-boxes system? First we can tabulate the macrostates and their multiplicity:

where the normalization constant (partition function) is just the sum of all the multiplicity:

$$Z = e^0 \cdot 1 + e^{-\frac{\epsilon}{kT}} \cdot 4 + e^{-\frac{2\epsilon}{kT}} \cdot 6 + e^{-\frac{3\epsilon}{kT}} \cdot 4 + e^{-\frac{4\epsilon}{kT}} \cdot 1$$

---

3 Speed Distribution of Ideal Gas

Because the energy of an ideal gas particle can be chosen from a continuous range of value, we should use the continuous version of Boltzmann distribution: writing the probability of a particle acquiring an energy in the range $[U_0, U_0 + \mathrm{d}U]$ as

$$P(U_0 < U < U_0 + \mathrm{d}U) = f(U_0)\,\mathrm{d}U = \frac{e^{-\frac{U_0}{kT}} \cdot W_p(U_0)}{Z}\,\mathrm{d}U$$

What is the function of multiplicity, $W_p(U)\,\mathrm{d}U$, for ideal gas particles?

3.1 Multiplicity of Ideal Gas

In the mechanics picture, microstates of a particle are described by its position and velocity:

$$\text{microstate} = (\vec{r}, \vec{v}) = (x, y, z, v_x, v_y, v_z)$$

Suppose the gas particle is in a container of size $L \times L \times L$ and there is no potential energy difference anywhere. The possibilities of the 6 state parameters are

- $x, y, z = $ Equal probability to be any real number within $[0, L]$.
- $v_x, v_y, v_z = $ Equal probability to be any real number as long as $\sqrt{v_x^2 + v_y^2 + v_z^2} < $ (speed of light)

To visualize, the possible states can be plotted in a 6-dimension state space of (x, y, z, v_x, v_y, v_z)

(add figure here: 3D x3D space; xyz form cube; v form sphere)

Because there is no potential energy difference anywhere in the box, we can choose the potential energy of the particle to be 0. So the particle's internal energy is equal to its kinetic energy, which is a function of the velocities only.

$$U = \text{K.E.} + \text{P.E.} = \frac{1}{2}m(v_x^2 + v_y^2 + v_z^2) + 0$$

So what is the multiplicity of a marcostate with internal energy in the range $[U_0, U_0 + \mathrm{d}U]$? It is the number of microstates (x, y, z, v_x, v_y, v_z) that satisfy $U_0 \leq \frac{1}{2}m(v_x^2 + v_y^2 + v_z^2) \leq U_0 + \mathrm{d}U$.

(add figure here: xyz every point satisfy anyway, v form sphere)

In the x-y-z space, every point is allowed.

In the v_x-v_y-v_z space, only the points with $|\vec{v}| = \sqrt{v_x^2 + v_y^2 + v_z^2}$ in between $\left[\sqrt{\frac{2U_0}{m}}, \sqrt{\frac{2(U_0+\mathrm{d}U)}{m}}\right]$ are allowed.

- Because multiplicity due to different (x, y, z) is the the same for any value of U, it shall be canceled by normalization when we compute probability. We can ignore it from now on.

- For the multiplicity due to different (v_x, v_y, v_z), although technically there are infinitely many points inside the spherical shell, this number of points must be proportional to the volume of the spherical shell.

$$(\text{Volume of the shell}) = 4\pi|\vec{v}|^2\,\mathrm{d}|\vec{v}| \quad \Rightarrow \quad (\text{\# of points within}) \propto 4\pi|\vec{v}|^2\,\mathrm{d}|\vec{v}|$$

Again, because any proportionality constant should be the same for any value of U, it shall be canceled by normalization when we compute probability.

By change of variable $|\vec{v}| = \sqrt{\frac{2U}{m}}$ and $\mathrm{d}|\vec{v}| = \sqrt{\frac{1}{2mU}}\,\mathrm{d}U$, we get $W(U)\,\mathrm{d}U$:

$$W(U_0)\,\mathrm{d}U = \begin{pmatrix} \text{\# of microstates that satisfy} \\ U \in [U_0, U_0 + \mathrm{d}U] \end{pmatrix}$$

$$= \begin{pmatrix} \text{\# of microstates that satisfy} \\ |\vec{v}| \in \left[\sqrt{\frac{2U_0}{m}}, \sqrt{\frac{2(U_0+\mathrm{d}U)}{m}}\right] \end{pmatrix}$$

$$\underset{\propto}{\propto} \begin{pmatrix} \text{Volume of spherical shell in between radius} \\ \sqrt{\frac{2U_0}{m}} \leq |\vec{v}| \leq \sqrt{\frac{2(U_0+\mathrm{d}U)}{m}} \end{pmatrix}$$

$$= 4\pi|\vec{v}|^2\,\mathrm{d}|\vec{v}|$$

$$= 4\pi\left(\frac{2U_0}{m}\right)\sqrt{\frac{1}{2mU_0}}\,\mathrm{d}U$$

$$= 4\pi\sqrt{\frac{2}{m^3}}\sqrt{U_0}\,\mathrm{d}U$$

Side note:

Note that because of the proportionality symbol above, the $W(U)$ we arrived at is NOT the number of state of ideal gas, but the **density of state** - the number of state per unit volume in the v_x-v_y-v_z space.

Calculating the exact number of state requires quantum mechanics. Because at the smallest scale, quantum mechanics tells us that internal energy of particles can only take quantized values. Then we can actually count the number of allowed states.

3.2 Normalization: Gaussian Integral

Now we can write down the Boltzmann distribution of ideal gas:

$$P(U) \sim f(U)\,\mathrm{d}U = \frac{e^{-\frac{U}{kT}} \cdot W_p(U)}{Z} = \frac{\overset{\text{constant}}{\overbrace{4\pi\sqrt{\frac{2}{m^3}}}} \cdot e^{-\frac{U}{kT}}\sqrt{U}\,\mathrm{d}U}{\cancel{Z}} = \frac{e^{-\frac{U}{kT}}\sqrt{U}\,\mathrm{d}U}{Z}$$

The remaining is to calculate the normalization constant Z. Let $x^2 = \frac{U}{kT}$ and $2x\,\mathrm{d}x = \frac{1}{kT}\,\mathrm{d}U$,

$$\underline{\int_0^\infty f(U)\,\mathrm{d}U} = \frac{1}{Z}\int_0^\infty e^{-\frac{U}{kT}}\sqrt{U}\,\mathrm{d}U$$

U of ideal gas = K.E. $\longrightarrow$ can only take positive values

$$= \frac{1}{Z}\int_0^\infty e^{-x^2}\sqrt{x^2 kT} \cdot (kT) \cdot 2x\,\mathrm{d}x$$

$$= \frac{2(kT)^{\frac{3}{2}}}{Z}\int_0^\infty e^{-x^2}x^2\,\mathrm{d}x$$

$$\equiv 1 \quad \text{(Required by normalization)}$$

This integral is extremely tricky to compute. First we do integration by-part,

$$\int_0^\infty e^{-x^2}x^2\,\mathrm{d}x = \int_0^\infty \frac{1}{2}e^{-x^2}x\,\mathrm{d}x^2$$

$$= -\int_0^\infty \frac{1}{2}x\,\mathrm{d}e^{-x^2}$$

$$= -\cancel{\left(\frac{1}{2}xe^{-x^2}\right)\Big|_0^\infty}^{\,\color{red}0} + \frac{1}{2}\int_0^\infty e^{-x^2}\,\mathrm{d}x$$

$$= \frac{1}{2}\int_0^\infty e^{-x^2}\,\mathrm{d}x$$

The integral on e^{-x^2} is called the **Gaussian integral**, which can be analytically computed only if the bounds are $\pm\infty$.

$$\boxed{\int_{-\infty}^\infty e^{-x^2}\,\mathrm{d}x \equiv \sqrt{\pi}}$$

But the range in our integral is only $[0, \infty]$, not $[-\infty, \infty]$, we only take half of it (because e^{-x^2} is an even function, $f(x) = f(-x)$). So the normalization constant Z is

$$Z = 2(kT)^{\frac{3}{2}}\int_0^\infty e^{-x^2}x^2\,\mathrm{d}x$$

$$= 2(kT)^{\frac{3}{2}} \cdot \frac{1}{2} \cdot \frac{\sqrt{\pi}}{2}$$

$$= \frac{\sqrt{\pi k^3 T^3}}{2}$$

The complete expression of Boltzmann distribution for ideal gas is therefore

$$f(U)\,\mathrm{d}U = \frac{2}{\sqrt{\pi k^3 T^3}} e^{-\frac{U}{kT}} \sqrt{U}\,\mathrm{d}U$$

Side note:

The standard method of computing Gaussian integral is as follow:

1. Let $I = \displaystyle\int_{-\infty}^{\infty} e^{-\frac{1}{2}ax^2}\,\mathrm{d}x$. The trick of computation is to promote it to 2D:

$$I^2 = \left(\int_{-\infty}^{\infty} e^{-\frac{1}{2}ax^2}\,\mathrm{d}x\right) \cdot \left(\int_{-\infty}^{\infty} e^{-\frac{1}{2}ay^2}\,\mathrm{d}y\right)$$

$$= \int_{x=-\infty}^{x=\infty} \int_{y=-\infty}^{y=\infty} e^{-\frac{1}{2}ax^2} e^{-\frac{1}{2}ay^2}\,\mathrm{d}y\,\mathrm{d}x$$

$$= \int_{x=-\infty}^{x=\infty} \int_{y=-\infty}^{y=\infty} e^{-\frac{1}{2}a(x^2+y^2)}\,\mathrm{d}y\,\mathrm{d}x$$

2. The we can do a change of coordinate from x-y to polar r-θ:

$$\iint f(x,y)\,\mathrm{d}x\,\mathrm{d}y \quad \Leftrightarrow \quad \iint f(r,\theta)\,r\,\mathrm{d}r\,\mathrm{d}\theta$$

- Converting the integrand:

$$e^{-\frac{1}{2}a(x^2+y^2)} \quad \Leftrightarrow e^{-\frac{1}{2}ar^2}$$

- The integration range is always the whole 2D plane:

$$(x,y) \in \{-\infty,\infty\} \times \{-\infty,\infty\} \quad \Leftrightarrow \quad (r,\theta) \in \{0,\infty\} \times \{0,2\pi\}$$

3. So the integral becomes

$$I^2 = \int_{r=0}^{r=\infty} \int_{\theta=0}^{\theta=2\pi} e^{-\frac{1}{2}ar^2}\,r\,\mathrm{d}\theta\,\mathrm{d}r$$

$$= 2\pi \cdot \int_{r=0}^{r=\infty} e^{-\frac{1}{2}ar^2}\,\mathrm{d}\left(\frac{r^2}{2}\right)$$

$$= \frac{2\pi}{a}$$

$$I = \sqrt{\frac{2\pi}{a}}$$

3.3 Speed Distribution & Average Speed

The final derivation is the famous Maxwell-Boltzmann distribution, which plots the probability of finding ideal gas particles at different speed v. The derivation is straightforward, by substituting $U = \frac{1}{2}mv^2$ and $\mathrm{d}U = mv\,\mathrm{d}v$ into the Boltzmann distribution.

$$f(U)\,\mathrm{d}U = \frac{2}{\sqrt{\pi k^3 T^3}} e^{-\frac{U}{kT}} \sqrt{U}\,\mathrm{d}U$$

$$= \frac{2}{\sqrt{\pi k^3 T^3}} e^{-\frac{mv^2}{2kT}} \sqrt{\frac{1}{2}mv^2} \cdot mv\,\mathrm{d}v$$

$$\boxed{f(v)\,\mathrm{d}v = \sqrt{\frac{2m^3}{\pi k^3 T^3}}\, v^2 e^{-\frac{mv^2}{2kT}}\,\mathrm{d}v}$$

In physics textbooks you can find 2 different formulas related to the average speed of ideal gas. these are derived as the expected values over the Maxwell-Boltzmann distribution.

1. **Average Speed** = Expected value to v, i.e. $\mathrm{E}[v]$.

$$\mathrm{E}[v] = \int_0^\infty v \cdot f(v)\,\mathrm{d}v$$

$$= \int_0^\infty v \cdot \sqrt{\frac{2m^3}{\pi k^2 T^3}}\, v^2 e^{-\frac{mv^2}{2kT}}\,\mathrm{d}v$$

$$= \sqrt{\frac{2m^3}{\pi k^2 T^3}} \left(\frac{2kT}{m}\right)^2 \int_0^\infty \underline{\left(\sqrt{\frac{m}{2kT}}v\right)^3 e^{-\left(\sqrt{\frac{m}{2kT}}v\right)^2} \mathrm{d}\left(\sqrt{\frac{m}{2kT}}v\right)}$$

$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad \uparrow$$

$$\text{\color{blue} Let } x = \sqrt{\tfrac{m}{2kT}}v$$

$$= \sqrt{\frac{32kT}{\pi m}} \underline{\int_0^\infty x^3 e^{-x^2}\,\mathrm{d}x}$$

$$\qquad\qquad \uparrow \underline{\quad} \text{\color{red} This integral can be computed}$$
$$\text{\color{red} by substitution and by-part}$$
$$\text{\color{red} No need to use Gaussian integral}$$

$$= \sqrt{\frac{32kT}{\pi m}} \cdot \frac{1}{2}$$

$$\boxed{v_{\mathrm{avg}} \overset{\mathrm{def}}{=} \mathrm{E}[v] = \sqrt{\frac{8kT}{\pi m}}}$$

2. **Root Mean Square Speed** $= \sqrt{\text{Expected value to } v^2}$, i.e. $\sqrt{\mathrm{E}[v^2]}$.

$$\mathrm{E}[v^2] = \int_0^\infty v^2 \cdot f(v)\,\mathrm{d}v$$

$$= \int_0^\infty v^2 \cdot \sqrt{\frac{2m^3}{\pi k^2 T^3}} v^2 e^{-\frac{mv^2}{2kT}}\,\mathrm{d}v$$

$$= \sqrt{\frac{2m^3}{\pi k^2 T^3}} \left(\frac{2kT}{m}\right)^3 \int_0^\infty \underline{\left(\sqrt{\frac{m}{2kT}}v\right)^3 e^{-\left(\sqrt{\frac{m}{2kT}}v\right)^2} \mathrm{d}\left(\sqrt{\frac{m}{2kT}}v\right)}$$

↑
Let $x = \sqrt{\frac{m}{2kT}}v$

$$= \sqrt{\frac{64k^2T^2}{\pi m^2}} \underline{\int_0^\infty x^4 e^{-x^2}\,\mathrm{d}x}$$

↑⌐ This requires computing
Gaussian integral

$$= \sqrt{\frac{64k^2T^2}{\pi m^2}} \cdot \frac{3\sqrt{\pi}}{8}$$

$$= \frac{3kT}{m}$$

$$\boxed{v_{\text{rms}} \overset{\text{def}}{=} \sqrt{\mathrm{E}[v^2]} = \sqrt{\frac{3kT}{m}}}$$

Because the average KE (which is something we can measure) is calculated as

$$\text{Avg. KE} = \mathrm{E}\left[\left(\frac{1}{2}m|v|^2\right)\right]$$

$$= \frac{1}{2}m\mathrm{E}\left[|v|^2\right]$$

$$= \frac{1}{2}mv_{\text{rms}}^2 = \frac{3}{2}kT$$

The root mean square speed is used much more commonly than the average speed.

— The End —