Introduction to Statistical Mechanics

by Tony Shing

Overview:

- Terminologies - Microstate, marcostate & multiplicity

- Boltzmann entropy

- Temperature scale by entropy

  - How to scale temperature for non ideal gas system?
  - What is actually negative temperature?

---

1 Microstate, Marcostate, Multiplicity

To begin with, we first emphasize the differences in state descriptions between the thermodynamics and mechanics.

1.1 Mechanics Picture: Microstates

In classical mechanics, the state of object can be described and predicted exactly. If we know the initial position x_0 and velocity v_0 of an object at the moment, then we can predict its future motion $(x(t), v(t))$ by solving differential equations (i.e. Newton's 2nd law).

(add figure here: some trajectory)

The procedure is the same for a system of many-body system. E.g. With 1000 objects, their motions can be predicted if we can measure the initial positions and velocities of all 1000 objects, and then solve a system of 1000 of 2nd order differential equations.

So the system's state can be uniquely described by 1000 initial positions and 1000 initial velocities. This is an example of **microstate description** - the state parameters are properties of individual objects.

$$\text{state} = (\vec{\boldsymbol{x}}_1, \vec{\boldsymbol{v}}_1, \vec{\boldsymbol{x}}_2, \vec{\boldsymbol{v}}_2, ..., \vec{\boldsymbol{x}}_N, \vec{\boldsymbol{v}}_N)$$

However taking measurement to all objects is impossible in reality. As we know that a few ten grams of matter already consist of $\sim 10^{23}$ atoms, mechanics picture is not suitable to model large amount of objects.

1.2 Thermodynamics Picture: Macrostates

In thermodynamics, state description is related to overall statistics of the system. E.g. We have been using the state parameters (P, V, T), which are all statistical quantities:

- $P \propto \underline{\text{Average}}$ collision force with the container.

- $V \propto \underline{\text{Average}}$ volume occupied by the gas.

- $T \propto \underline{\text{Average}}$ velocity of particles.

(add figure here: vol, temp)

This is the **macrostate description** - states are described by the objects' collective behavior instead of looking into them one by one. Description by macrostate is much simpler than by microstate because it requires much fewer number of parameters (from $\sim 10^{23}$ to around 10). However this also implies a lost of details about the system.

1.3   Multiplicity

Microstates and marcostates are in a many-to-one correspondance:

- It we can measure all the microstate parameters, we can calculate the statistical quantities and determine which marcrostate the system is in.

- It is impossible to determine the microstate by measuring macrostate parameters, because macrostate description doesn't give fine details. We can have many different configurations (microstates) that yield the same statistical properties (macrostate).

The **multiplicity W of a macrostate** is defined as the number of microstates that correspond to it.

(add figure here: box of gas)

Example 1.1. Suppose we have 4 balls that are allowed to move between 2 boxes L and R. What are the

1. Microstate description?

2. Macrostate description?

3. Multiplicity of each macrostate?

(add figure here: ball in box)

1. For microstate, we have to look at the configuration of the individual balls - for each ball whether it is in the L box or the R box. There are $2^4 = 16$ different microstates in total.

2. For macrostate, we look at collective behavior of the balls - how many balls are in the L and R box respectively. There are 5 different kinds of macrostates in total.

3. By binomial theorem, The multiplicity of each of the macrostates are $C_L^4 = \{1, 4, 6, 4, 1\}$ respectively.

2 Boltzmann Entropy

2.1 Origin of Irreversibility

Imagine we allow the balls to move freely between the two boxes. *Assume* the two boxes to be identical (Can this assumption be wrong?), so that each ball has equal probability to be found in either box. Then

– Each microstate has an equal probability $\frac{1}{16}$ to be found.

– The probability of being in a macrostate is proportional to its multiplicity.

Therefore, if we take many snapshot from time to time, we will find that the macrostate 2L-2R has the highest probability $\frac{6}{16}$ to be observed.

<p style="color:red; text-align:center;">(add figure here: box)</p>

Now consider the a more extreme case, say we begin with 1000 balls in the L box. After allowing the balls to move freely between the boxes, what will we observe?

– The multiplicity of macrostate 1000L-0R is $C_0^{1000} = 1$, so its probability of occurence is $\frac{1}{2^{1000}} \sim 10^{-301}$.

– The multiplicity of macrostate 500L-500R is C_{500}^{1000}, so its probability of occurence is $\frac{C_{500}^{1000}}{2^{1000}} \sim 10^{-2}$.

Which means on average, we will observe the balls evenly distributed two boxes for 10^{299} times before observing them all on 1 side. But our universe is only 13 billion years old $\sim 10^{14}$ seconds - This is equivenlently saying that 1000L-0R state could never be observed again.

– We can always see the system evolves from 1000L-0R to 500L-500R.

– We can almost never see the system evolves from 500L-500R to 1000L-0R.

<p style="color:red; text-align:center;">This is the irreversibility!</p>

2.2 Boltzmann's Hypothesis

Summarizing what we have learnt:

– Clausius found a function S, which the change $\Delta S \geq 0$ in every closed system (energy conserved within).

– According to probability, system always evolves from marcostate with low multiplicity to macrostate with high multiplicity. The reverse (almost) never happens.

So a smart guy, Ludwig Boltzmann proposed the relation:

$$S = \phi(W)$$

Entropy   Some relation   Multiplicity
           to be determined

What mathematical requirements should $\phi(\cdot)$ satisfy? Consider a closed system of 2 boxes A and B, which are in a macrostate with a multiplicity W_A and W_B respectively.

   – If we look at them independently,

$$S_{\text{total}} = S_A + S_B = \phi(W_A) + \phi(W_B)$$

(add figure here: separate box)

   – If we look at them as a combined system, because calculating multiplicity is just counting the number of combination,

$$W_{\text{total}} = W_A \times W_B$$
$$\Rightarrow \quad S_{\text{total}} = \phi(W_A \times W_B)$$

What function(s) satisfy such relation $\phi(W_A) + \phi(W_B) = \phi(W_A \times W_B)$? **Logarithm**.

$$\ln(W_A) + \ln(W_B) = \ln(W_A \times W_B)$$

This is how we arrive at the Boltzmann entropy formula:

$$\boxed{S = k \ln W}$$

k = Some proportionality constant we don't know yet

2.3   Finding Boltzmann Constant

Now it comes to the question of what value should we take for k. Ideal gas, the system being most familiar by physicists, is used to set the baseline.

1. Firstly, we consider the gas expanding under an isothermal process. The entropy change is

$$\Delta S = \int \frac{\delta Q}{T} = nR \ln \left(\frac{V_2}{V_1} \right)$$

2. Change in multiplicity on an isotheral path can be modelled by

   i Divide the box volume V_1 into grids of (arbituary) size ΔV.

(add figure here: divide grid)

   ii Put n particles in the box. Then

   – # of ways to put 1 particle $= \dfrac{V_1}{\Delta V}$

P.4

- – # of ways to put N particles $= \left(\dfrac{V_1}{\Delta V}\right)^N$

So the multiplicity is proportional to $W_1 \propto \left(\dfrac{V_1}{\Delta V}\right)^N$.

iii Let the box expand to volume V_2. Now there are more grids in the box.

- – # of ways to put 1 particle $= \dfrac{V_2}{\Delta V}$

- – # of ways to put N particles $= \left(\dfrac{V_2}{\Delta V}\right)^N$

So the multiplicity increases to $W_2 \propto \left(\dfrac{V_2}{\Delta V}\right)^N$.

3. Plugging in the above into the formula $S = \ln(W)$,

$$\Delta S = S_2 - S_1$$

$$nR\ln\left(\frac{V_2}{V_1}\right) = k\ln(W_2) - k\ln(W_1)$$

$$= k\ln\left[\frac{\frac{V_2}{\Delta V}}{\frac{V_1}{\Delta V}}\right]$$

$$= Nk\ln\left(\frac{V_2}{V_1}\right)$$

$$\underline{nR} = \underline{Nk}$$

# of particle but in moles   ↑ Ideal gas constant   # of particle

$$k = \frac{R}{\underline{N_A}} \longleftarrow \text{Avogadro's number}$$

$$\approx 1.38 \times 10^{-23}$$

Example 2.1. We can use Boltzmann's entropy formula to estimate the change multiplicity in real life process, to see how much "irreversible" they are.

$$\Delta S = k\ln(\Delta W) \quad \Rightarrow \quad \Delta W = e^{\frac{\Delta S}{k}}$$

Suppose we have 2 objects with the same heat capacity $C = 1$, but object A is at 300.1 K and object B is at 299.9 K. After allowing heat exchange, we expect both reaching the equilibrium temperature 300 K. The entropy change in this process is

$$\begin{cases} \Delta S_A = \displaystyle\int_{300.1}^{300} \frac{C\,dT}{T} = C\cdot\ln\left(\frac{300}{300.1}\right) \\[2mm] \Delta S_B = \displaystyle\int_{299.9}^{300} \frac{C\,dT}{T} = C\cdot\ln\left(\frac{300}{299.9}\right) \end{cases}$$

$$\Rightarrow \quad \Delta S_{\text{total}} = C\cdot\ln\left(\frac{300}{300.1}\cdot\frac{300}{299.9}\right) \approx 1.11 \times 10^{-7}$$

$$\Delta W = \frac{W_{\text{final}}}{W_{\text{initial}}} = e^{\frac{\Delta S}{k}} = e^{8.05\times10^{15}} \approx 10^{10^{15.54}}$$

The multiplicity has increased by a ratio of a number with 10^{15} digits. This is physically equivenlently to say that the reverse process will never be observed.

2.4   Entropy as Disorder-ness

In physics textbook, we regards entropy as disorderness due to Boltzmann hypothesis. The idea is intuitive - there are more ways to make things look messy than to make things look tidy, so a messy state implies higher multiplicity and thus higher entropy. Just like clothes should be in wardrobes and books should be on bookshelves if you want your room looking clean, but you can put them anywhere else if you want your room looking messy.

---

3   A Temperature Scale by Entropy

Let's recap some temperature scales we know in history:

1. <u>Emperical scales</u>: A linear interpolation between two fixed reference point. E.g.

   - Celsius (1701-1744) : Between water's freezing point (0°C) and boiling point (100°C)
   - Fahrenheit (1686-1736) : Between the eutectic temperature of ammonium chloride brine (0°F) and average human body temperature (96°F)

   Thermometers in these scales measure temperature using material expansion under heat, which limit the scale by the freezing point of the measuring material.

2. <u>Thermodynamics scale</u>: A theoretical scale based entirely based on macroscopic properties of material. i.e. Kelvin scale based on relation with Carnot engine's efficiency.

   - The first scale that defines the full temperature range -273.15°C to ∞°C.
   - Prove that going below -273.15°C is not physical.

However after entropy is established as a state function, physicists find a new mathematical way to define temperature - through the relation between internal energy and entropy.

   - We previously write internal energy U of a material as a state function of 3 variables (P, V, T), but only two of them are independent because there is the equation of state (e.g. ideal gas law).

   - Entropy is also a function of (P, V, T).

It should not make a different if we do a change of variable, making internal energy as a function of 4 variables (P, V, T, S)!

$$U = U(P, V, T, S) \qquad \text{with} \qquad \text{only 2 of the variables independent}$$

Let's say if we choose the two independent ones to be V and S,

   - According to the thermal 1$^{\text{st}}$ law,

$$\mathrm{d}U = \delta Q - \delta W = T\,\mathrm{d}S - P\,\mathrm{d}V$$

– By partial differentiation chain rule,

$$dU(S,V) = \left(\frac{\partial U}{\partial S}\right) dS + \left(\frac{\partial U}{\partial V}\right) dV$$

T and P appears as functions of (S,V), giving us new way of defining them.

$$T(S,V) \overset{\text{def}}{=} \frac{\partial U}{\partial S} \quad , \quad P(S,V) \overset{\text{def}}{=} -\frac{\partial U}{\partial V}$$

If we can measure a material's internal energy and entropy, plot them on a graph, then temperature of the material can be defined using the slope of the U-S graph.

However, because U is usaually some simple function and easy to compute, while S needs to be computed by multiplicity (i.e. a lot of combinatorics), it is easier to define T by differentiating in the reverse way:

$$\boxed{\frac{1}{T} \overset{\text{def}}{=} \frac{\partial S}{\partial U}}$$

Notations:

In some advanced thermodynamics textbooks, the definition may be subscripted in the following way:

$$T(S,V) \overset{\text{def}}{=} \left(\frac{\partial U}{\partial S}\right)_{\text{const. V}} \quad , \quad P(S,V) \overset{\text{def}}{=} -\left(\frac{\partial U}{\partial V}\right)_{\text{const. S}}$$

This is just for emphasizing which of the (P,V,T,S) is the another independent variable, which must be kept constant in the partial differentiation.

By this temperatue - internal energy - entropy relation, we can derive some interesting properties about non-ideal gas materials.

1. Count multiplicity by some state parameter x. x is specific to the material. Then compute $S = k \ln W$

2. Express U as a function of x.

3. A relation is born by relate S and U by T.

$$\frac{1}{T} = \frac{\partial S}{\partial U} = \frac{\left(\dfrac{\partial S}{\partial x}\right)}{\left(\dfrac{\partial U}{\partial x}\right)}$$

3.1 Example 1: Crystal Vaccancy Model

A crystal is a regular arrangement of atoms/molecules.

Pure crystal

(add figure here: pure crystal)

There is only 1 possible arrangement. $\Rightarrow$ Multiplicity $= 1$.

Crystal with defect

(add figure here: Crystal with defect)

For example, vaccancy (missing atoms), or having impurities at some position. Have many combination. $\Rightarrow$ High multiplicity

In a crystal with defect, we also expect there is a energy change per defect because the atomic bondings around the defect are different. Thus we can use the temperature - internal energy - entropy to model the expected number of defects.

1. Suppose we have a crystal with a total $N + n$ sites, N being the crystal's atoms and n being substitute by impurity/vaccancy. Then the multiplicity is the number of ways to arrange N atoms into $N + n$ positions.

$$W = C_n^{N+n} = \frac{(N+n)!}{N!n!}$$

2. Calculating the Boltzmann entropy:

$$\begin{aligned}
S &= k \ln \left(\frac{(N+n)!}{N!n!} \right) \\
&= k[\ln(N+n)! - \ln N! - \ln n!] \\
&\approx k[(N+n)\ln(N+n) - N\ln N - n\ln n]
\end{aligned}$$

Here we have used the **Stirling approximation**: $\ln N! \approx N \ln N - N$, when N is a very big number.

3. *Assume* that each impurity/vaccancy brings an increase of internal energy by ϵ, which ϵ may be some values measurable from experiments. A total n sites of impurity/vaccancy gives a total internal energy

$$U = n\epsilon$$

4. Finally, apply the T-S-U relation. We observe that both S and U are functions of n,

$$\frac{\partial S}{\partial n} = k\frac{\partial}{\partial n}[(N+n)\ln(N+n) - N\ln N - n\ln n]$$
$$= k[\ln(N+n) - \ln n]$$
$$= k\ln\left(\frac{N+n}{n}\right)$$

$$\frac{\partial U}{\partial n} = \frac{\partial}{\partial n}n\epsilon = \epsilon$$

$$\Rightarrow \quad \frac{1}{T} = \frac{\partial S}{\partial U} = \frac{\left(\dfrac{\partial S}{\partial n}\right)}{\left(\dfrac{\partial U}{\partial n}\right)} = \frac{k}{\epsilon}\ln\left(\frac{N+n}{n}\right)$$

By a change of subject, we get

$$\text{(Portion of defects)} = \frac{n}{N} = \frac{1}{e^{-\frac{\epsilon}{kT}} - 1}$$

Observe that $n \to 0$ only if $T \to 0$, meaning that crystal in real life must have some defects (statistically). Crystal can spontaneously form defects under normal temperature.

3.2 Example 2: Rubber Band Contraction

A rubber band can be considered a long chain of atom (e.g. polymer).

Full extension

(add figure here: straight)

When being pulled straight, its length is the longest. There is only 1 possible arrangement. $\Rightarrow$ Multiplicity $= 1$.

Entangled state

(add figure here: entangled)

The string curled up and the total length shorten. Have many combination. $\Rightarrow$ High multiplicity

For simplicity, we only consider a 1D chain - the connections to next atom can only be pointing left or right.

(add figure here: 1D chain)

1. Suppose the chain has $N+1$ atoms (so a total of N connections). There are n_L connections pointing to the left and $N - n_L$ connections pointing to the right. Then the multiplicity is the number of ways to assign n_L left connections out of the total N connections.

$$W = C_{n_L}^N = \frac{N!}{n_L!(N - n_L)!}$$

2. Calculating the Boltzmann entropy:

$$S = k \ln \left(\frac{N!}{n_L!(N - n_L)!} \right)$$
$$= k[\ln N! - \ln n_L! - \ln(N - n_L)!]$$
$$\approx k[N \ln N - n_L \ln n_L - (N - n_L) \ln(N - n_L)]$$

Again, the Stirling approximation formula is used.

3. We can relate the internal energy in a rubber band with its stored elastic energy. Let the connection between atoms be of length d, such that the total length of the chain is $[n_L - (N - n_L)] \times d = (2n_L - N)d$. When a tension F is applied on both side of the chain to unfold 1 right connection into 1 left connection,

<div style="color:red; text-align:center">(add figure here: unfold chain)</div>

   − Number of left connection $n_L \to n_L + 1$.
   − Total length of the chain $(2n_L - N)d \to (2n_L - N)d + 2d$.
   − The work done by the tension is $F \cdot 2d$

Denote Δn as the number of right connection being unfolded into left connection, then

$$\frac{\partial n_L}{\partial(\Delta n)} \approx \frac{(n_L + \Delta n) - n_L}{\Delta n} = 1$$

$$\frac{\partial U}{\partial(\Delta n)} = \frac{\partial(W.D.)}{\partial(\Delta n)} \approx \frac{(F + 2Fd(\Delta n)) - F}{\Delta n} = 2Fd$$

4. Finally, apply the T-S-U relation. We observe that both S and U are functions of Δn

$$\frac{\partial S}{\partial(\Delta n)} = \frac{dS}{dn_L} \frac{\partial n_L}{\partial(\Delta n)}$$
$$= k[\ln(N - n_L) - \ln n_L] \cdot 1$$
$$= k \ln \left(\frac{N - n_L}{n_L} \right)$$

$$\frac{\partial U}{\partial(\Delta n)} = 2Fd$$

$$\Rightarrow \quad \frac{1}{T} = \frac{\partial S}{\partial U} = \frac{\left(\frac{\partial S}{\partial n} \right)}{\left(\frac{\partial U}{\partial n} \right)} = \frac{k}{2Fd} \ln \left(\frac{N - n_L}{n_L} \right)$$

Substitute by the length of the chain $L = (2n_L - N)d$ and change of subject to F, we get

$$(\text{Tension}) = F = \frac{kT}{2d} \ln \left(\frac{Nd + L}{Nd - L} \right)$$

Observe that F increases if T increases. <u>The chain of atoms would like to contract under heat!</u>

3.3   Negative Temperature?

In our new definition of temperature

$$\frac{1}{T} = \frac{\partial S}{\partial U} = \left(\begin{matrix}\text{Rate of change of entropy} \\ \text{with respect to energy}\end{matrix}\right)$$

if there is a system whose entropy decreases ($S \downarrow$) when its internal energy increases ($U \uparrow$), $\frac{1}{T} = \frac{\partial S}{\partial U} < 0$, i.e. negative temperature.

Negative temperature happens only in systems with finite number of microstate.

For example, let's revisite the 2 boxes system. But this time we assign potential energy difference between the boxes - balls in the right box would have higher energy than in the left box by ϵ.

(add figure here: energy level)

When total energy $= 3\epsilon$, there are 4 possible microstates.

(add figure here: 4 microstate)

When total energy $= 4\epsilon$, there is only 1 possible microstates.

(add figure here: 1 microstate)

(add figure here: W vs E)

Imagine there are many more particles, the bar chart will become a smooth curve like so:

(add figure here: smooth curve)

The temperature scale defined by entropy is kind of "weird"...

(add figure here: low energy -¿ high energy)

In this sense, temperature is no long about hot or cold anymore.

— The End —