

# Entropy Demystified Note

Yang Xue

March 2024

## 1 Mathematical Preparation

We will introduce some basic concept in calculus. It won't be rigorous but we will provide some example which should be enough to actually do calculation for entropy.

**Definiton 1** *If the deviation of the value  $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  from  $f(x)$  evaluated at some point  $x'$  in some neighborhood of  $x$  can be approximated linearly as weighted sum of difference in each component  $dx_i = x'_i - x_i$ , or more succinctly,*

$$df = A_1(x)dx_1 + \dots A_n(x)dx_n$$

where  $df$  is the approximated deviation of value from  $x$ , then we say  $df$  is differentiable at  $x$ . This approximation business looks sketchy but there is a mathematically rigorous definition that define what counts as a good approximation and there is at most one (there could be none) such approximation at any point.

Suppose now we only have one input, i.e.  $f : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ , then it's easy to see what's this best linear approximation, it's the rise over run, the derivative of  $f$  at  $x$ . But now we have multiple inputs, calculus tells us we can simply calculate the contribution from each the difference in direction  $dx_i$  as if other directions don't exist.

**Proposition 1** *If  $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  is differentiable, then  $A_i(x) = \frac{\partial f(x)}{\partial x_i}$ , where  $A_i$  is defined in definition 1. The notation  $\partial$  is a fancy way of saying, I only care about  $i^{\text{th}}$  component and other variable will be treated as parameters.*

**Example 1** *Let's first review a one dimension case. What's the derivative of  $f(x) = x^2$ ?*

$$\frac{df(x)}{dx} := \lim_{\Delta x \rightarrow 0} \frac{(x + \Delta x)^2 - x^2}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{2x\Delta x + (\Delta x)^2}{\Delta x} = 2x$$

Now let's introduce one more variable  $f(x, y) = x^2y$ . We have

$$\frac{\partial f(x, y)}{\partial x} = \lim_{\Delta x \rightarrow 0} \frac{(x + \Delta x)^2y - x^2y}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{2x\Delta xy + (\Delta xy)^2}{\Delta x} = 2xy$$

$$\frac{\partial f(x, y)}{\partial y} = \lim_{\Delta y \rightarrow 0} \frac{x^2(y + \Delta y) - x^2y}{\Delta y} = x^2$$

Hence the best linear approximation at any point  $x$  is

$$df = 2xydx + x^2dy$$

## 2 What is entropy?

### 2.1 Classical Thermodynamic

**Definiton 2** *In thermodynamics we don't care about the movement of individual particles but the collective behavior of them. A macrostate is then specified with the macroscopic property like temperature, volume and number of particles.*

**Definiton 3** Entropy is defined up to a constant. The change in entropy of a system during two state is

$$dS = \frac{dQ}{T}$$

where  $dQ$  is the heat that goes into the system,  $T$  is the temperature of the system, both of which are calculated during a reversible process.

**Remark 1** Without going in cycle on what is a reversible process, in current situation, a reversible process is simply a process that can be reversed, i.e. I can return the universe back to what's before the process. An example of a reversible process would be slowly pushing the a perfect piston of a sealed chamber; an irreversible process would be hot and cold stuffs kept in contact with each other and reaching the same temperature in the end.

Now does that mean we don't have a formula for entropy in a irreversible process? The answer is no. We can always connect two macrostates with a reversible process, even though in reality what happens is a irreversible process.

It can be seen as a fundamental assumption that the change in entropy of the system is irrelevant to how we connect two states, as long as we connect them with an reversible process. This allows us the have a state function  $S$  of macroscopic property like internal energy  $U$ , volume  $V$ , number of particles  $N$ , because we have a well-defined way to calculate the difference in two states that only depends on the initial and final  $U$ ,  $V$ ,  $N$ , etc.

Now we have the language for following laws.

**Law of Physics 1** First Law of Thermodynamics: In a closed system,

$$dU = dW + dQ$$

where  $W$  is work done on the system and  $Q$  is heat transferred into the system. A closed system is a system where there's no exchange of energy or particles.

Second Law of Thermodynamics: The entropy of a closed system is non-decreasing.

This is all cute. Let's put these into work

**Example 2** Ideal gas is defined to be the gas that satisfies  $PV = Nk_B T$ ,  $U = Nk_B T$ , where  $P$  is pressure,  $V$  is volume,  $N$  is number of particles  $T$  is temperature and  $k_B$  is a constant. Suppose now we have a box of gas with volume  $V_1$  with a piston, in contact with a heat bath with temperature  $T$ . Now we pull the piston so the gas expands. Let's further suppose the expansion is slow enough so it stays at the temperature  $T$ . The process ends with the volume  $V_2$ . What's the entropy change?

Since

$$dS = \frac{dQ}{T}$$

and temperature is constant by assumption

$$\Delta S = \frac{\Delta Q}{T}$$

By first law of thermodynamics,

$$\Delta Q = \Delta U - \Delta W$$

where  $\Delta W$  is the work done on the system, which can be calculate by

$$dW = -PdV = -\frac{Nk_B T}{V}dV$$

which, you must believe me, gives

$$\Delta W = -Nk_B T \ln\left(\frac{V_2}{V_1}\right)$$

So the heat exchange is

$$\Delta Q = Nk_B T \ln\left(\frac{V_2}{V_1}\right)$$

Note the sign is positive. That means the system absorbs heat, and thus the entropy increases, since there's no change in temperature and by assumption no change in internal energy. The entropy change is

$$\Delta S = Nk_B \ln\left(\frac{V_2}{V_1}\right)$$

This is all good but consider now a similar process. We have a insulated box partitioned into two, separated by a plank. We fill the same gas in one of the partition make the other vacuum. We remove the partition suddenly. The gas diffuses with no change in temperature (because no work has been done). What's the entropy change? Is it 0 in this case? No! it's not a reversible process, even though it's the process that actually happened. We need to choose the reversible process, one example would be our first process, so the entropy change should be the same.

## 2.2 Statistical Mechanics

When I was learning classical thermodynamics, I found all of these very confusing and complicated. Thankfully physicists figured out another formalism which makes more sense to me.

**Definiton 4** A microstate is specified by the individual state of particles. Entropy is proportional to the logarithm of multiplicity of a macrostate. Multiplicity  $\Omega$  of a macrostate is the number of microstates that gives rise to the macrostate. That is

$$S = k_B \ln(\Omega)$$

**Remark 2** A macrostate is usually described by  $U, V, N$ , but as we described before, they don't tell us how individual particle moves. Now why taking the log? This is because we want entropy of two system to be sum of the individual systems not multiplication. The  $k_B$  is called Boltzmann constant. It's there purely for historic reason. We want to make two definitions of entropy agree.

**Law of Physics 2** This is not a law but more like physicists' wish. The time a system with energy  $E$  spends being in any accessible microstate is the same. Here "accessible" simply emphasizes the energy of the microstate is  $E$ . This assumption is called ergodic hypothesis. Almost all system studied in physics can be proved to have this property.

With this hypothesis, second law of thermodynamics is tautological. Basically, it's saying the closed system is most likely to go to the macrostate with more microstates. Why? Because in most systems, that means they are most probable.

**Example 3** Now revisit our previous example. It's still not clear with a classical system like this, what do we mean by the number of microstates? There seems to be infinitely many for any macrostate. The precise meaning of this requires us to go to quantum mechanics. But if we are allowed to wave our hand, we might simply argue when the size of the box goes from  $V_1$  to  $V_2$ , the number of microstates also grows by the same ratio, simply because we have more locations to choose for each particle. Denote multiplicity by  $\Omega$ , we have

$$\Delta S = S_2 - S_1 = k_B \ln(\Omega_2) - k_B \ln(\Omega_1) = k_B \ln\left(\frac{\Omega_2}{\Omega_1}\right)$$

We this powerful new definition of entropy in hand we can actually redefine temperature! Why do we want to do that? As it turns out for a long time we define temperature experimentally without knowing what it is really. That is, temperature is whatever the thermometer says.

**Definiton 5** Let  $T$  be the temperatue of the system,  $S$  entropy,  $U$  energy,  $V$  volume,  $N$  number of particles.

$$\frac{1}{T} = \frac{\partial S(U, V, N)}{\partial U}$$

Why should we trust this is the correct definition? Now suppose two systems of temperature  $T_1$  and  $T_2$  are in contact with each other but otherwise can not exchange energy with environment. Their temperatures should be the same in the end with our experimental definition of temperature. Are By the second law

$$S(U_1, V_1, V_2, N_1, N_2) = S_1(U_1, V_1, N_1) + S_2(U - U_1, V_2, N_2)$$

is maximized. The assumption that two systems are insulated forces us to constrain  $U_2 = U - U_1$ . This entails

$$dS = \frac{\partial S_1}{\partial U_1} dU_1 - \frac{\partial S_2}{\partial U_1} dU_1 + \frac{\partial S_1}{\partial V_1} dV_1 + \frac{\partial S_1}{\partial N_1} dN_1 + \frac{\partial S_2}{\partial V_2} dV_2 + \frac{\partial S_2}{\partial N_2} dN_2 = 0$$

because otherwise there's some direction you can go that can make entropy larger. Hence

$$\frac{\partial S_1}{\partial U_1} = \frac{\partial S_2}{\partial U_2}$$

$$T_1 = T_2$$

If we put constraint on total volume and total particle we would get definition of pressure and chemical potential but in summary.

$$dS = \frac{1}{T} dU + \frac{P}{T} dV - \frac{\mu}{T} dN$$

This is called Thermodynamic Identity. This is cute. But can we actually calculate something with this? Sadly the calculation for ideal gas is quite convoluted. However it's somewhat straight forward to calculate this model called Einstein Solid.

**Example 4** *In Einstein solid, each atom is thought to be a quantum harmonic oscillator, the quantum equivalent of ball on a spring. Due to results from quantum mechanics, we find it's energy when measured can only take on*

$$E_n = \hbar\omega\left(\frac{1}{2} + n\right)$$

*For our purpose, the constant energy shift is unimportant, it just make calculation awkward so let's drop it for now. Suppose our system has  $N$  harmonic oscillator ( $\frac{N}{3}$  atoms), what's the entropy? We start by calculate the multiplicity when the system has in total energy  $x\hbar\omega$ , or in other words  $x$  quanta of energy. How many ways I can distribute the energy among the atoms? Imagine these quanta stands in line waiting for the oscillators to take them. The atoms can pick arbitrary numbers of quanta. This is equivalent to dividing the quanta with bars, each segment are the quanta taken by an atom. The first segment is for the first oscillator, the second segment is for the second oscillator and so on. Each distribution corresponds to one way of choosing  $x$  object to be quanta from  $N+x-1$  object. In the counting notation*

$$\Omega = \binom{N+x-1}{x} = \frac{(N+x-1)!}{x!(N-1)!}$$

So

$$S = k_B \ln\left(\frac{(N+x-1)!}{x!(N-1)!}\right) = k_B(\ln(N+x-1)! - \ln x! - \ln(N-1)!)$$

where  $x = \frac{U}{\hbar\omega}$ . In principle we are done, but really we want to make the result more manageable when manipulate the entropy to get other quantities. Now we introduce sterling approximation

$$\ln n! \approx n \ln n - n.$$

for large  $n$ . This allow us to write.

$$\begin{aligned} S &\approx k_B((N+x) \ln(N+x) - (N+x) - (x \ln(x) - x) - (N \ln N - N)) \\ &= k_B((N+x) \ln(N+x) - x \ln(x) - N \ln N) \\ &= k_B\left(\left(N + \frac{U}{\hbar\omega}\right) \ln\left(N + \frac{U}{\hbar\omega}\right) - \frac{U}{\hbar\omega} \ln\left(\frac{U}{\hbar\omega}\right) - N \ln N\right) \end{aligned}$$

where we have use the fact that  $N$  is large so we can forget about the 1. We can get the relation between temperature and energy (I won't bore you with the calculation)

$$\frac{1}{T} = \frac{\partial S}{\partial U} = k_B \frac{-\ln\left(\frac{U}{\hbar\omega}\right) + \ln\left(N + \frac{U}{\hbar\omega}\right)}{\hbar\omega} = \frac{k_B}{\hbar\omega} \ln\left(\frac{N\hbar\omega + U}{U}\right)$$

Invert the relation

$$U = \frac{N\hbar\omega}{e^{\frac{\hbar\omega}{k_B T}} - 1}$$

At high temperature  $\exp(\frac{\hbar\omega}{k_B T}) \approx 1 + \frac{\hbar\omega}{k_B T}$ , as you can check by noting  $1+x$  is the best linear approximation for  $e^x$  at  $x=0$ . Hence at high T,

$$U = Nk_B T$$

which is independent of Planks Constant. This agrees with experimental result called law of Dulong-Petit. In fact this is the manifestation of equipartition theorem.

**Definiton 6** *Hamiltonian of a system  $H(p_i, q_i)$  is a functional from real functions to  $\mathbb{R}^n$ , such that it gives the correct equation of motion through*

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}.$$

*Without going into too much detail,  $q_i$  describes the position of each the particle and  $p_i$  is the momentum of each particle, and Hamiltonian is the total energy. The rigorous treatment of this is a subject of analytical mechanics.*

**Theorem 1** *Suppose the Hamiltonian of the system  $H$  depends quadratically on  $N$  degrees of freedoms  $p_i$  and  $q_i$  then classically*

$$E(p_i, q_i) = \frac{N}{2} k_B T$$

**Remark 3** *This theorem is best proved in canonical ensemble. The Hamiltonian for Einstein solid is*

$$H = \sum_i \left( \frac{p_i^2}{2m} + \frac{1}{2} m \omega x_i^2 \right)$$

*There's 2 degree of freedom per harmonic oscillator, which gives our result at high temperature!*

### 3 Appendix: Technical Details of Math

**Definiton 7** *Let  $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ . Suppose for  $\forall \epsilon > 0, \exists \delta > 0$ , such that  $|f(x) - A| \leq \epsilon$  for  $|x - x_0| < \delta$ , then we say*

$$\lim_{x \rightarrow x_0} f(x) = A$$

**Definiton 8** *Let  $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ . Suppose*

$$f(x + \Delta x) = A_1(x)\Delta x_1 + A_2(x)\Delta x_2 \dots + A_n(x)\Delta x_n + o(\Delta x)$$

*where  $\Delta x_i$  is the  $i$ th component of vector  $\Delta x$  and*

$$\lim_{\Delta x \rightarrow 0} \frac{o(\Delta x)}{|\Delta x|} = 0,$$

*then we denote*

$$df = A_1(x)dx_1 + A_2(x)dx_2 + \dots + A_n(x)dx_n$$

**Remark 4** *We can roughly think "d" as an operator that extracts the linear part of the deviation  $\delta y$  when we move away from  $x$ .*

**Definiton 9** *Let  $f: \mathbb{R}^n \rightarrow \mathbb{R}$ , we define the partial derivative of  $f$  at  $x$  as*

$$\frac{\partial f}{\partial x_i}(x) = \lim_{\Delta x_i \rightarrow 0} \frac{f(x_1, \dots, x_i + \Delta x_i, \dots, x_n) - f(x_1, x_2, \dots, x_n)}{\Delta x_i}$$

*, if the limit exists. Specially, if  $n=1$ , then we use the notation*

$$\frac{df}{dx}(x) = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

**Remark 5** I wrote  $\frac{\partial f}{\partial x_i}(x)$  with  $x$  outside of the fraction to emphasize really taking partial is a map from function to a function which we then evaluate at  $x$ .

The notation for  $n=1$  is justified by following result.

**Proposition 2** Let  $f: \mathbb{R}^n \rightarrow \mathbb{R}$ . Suppose

$$df = A_1(x)dx_1 + A_2(x)dx_2 + \dots + A_n(x)dx_n,$$

then

$$A_i(x) = \frac{\partial f(x)}{\partial x_i}.$$

In case of  $n=1$

$$df = \frac{df(x)}{dx} dx$$

which looks tautological but really those  $d$ 's have different meaning.